TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Hari Nortunen

Light, Shape and Space

Direct and Inverse Problems of Individual Objects and Populations

Julkaisu 1555• Publication 1555

Tampere 2018

Tampereen teknillinen yliopisto. Julkaisu 1555
Tampere University of Technology. Publication 1555

Hari Nortunen

Light, Shape and Space

Direct and Inverse Problems of Individual Objects and Populations

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB214, at Tampere University of Technology, on the 29th of June 2018, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2018

Doctoral candidate: Hari Nortunen
Laboratory of Mathematics
Faculty of Natural Sciences
Tampere University of Technology
Finland

Supervisor: Mikko Kaasalainen, Professor
Laboratory of Mathematics
Faculty of Natural Sciences
Tampere University of Technology
Finland

Pre-examiners: Erkki Somersalo, Professor
Department of Mathematics, Applied Mathematics and Statistics
Case Western Reserve University
United States of America

Juha Vierinen, Associate Professor
Department of Physics and Technology
University of Tromsø – The Arctic University of Norway
Norway

Opponent: Tanja Tarvainen, Associate Professor
Department of Applied Physics
University of Eastern Finland
Finland

ISBN 978-952-15-4164-3 (printed)
ISBN 978-952-15-4169-8 (PDF)
ISSN 1459-2045

Abstract

In this thesis, we examine two case studies. The first case studies how the shape of a celestial body influences an effect caused by light, while the second case analyzes how the light, in the form of brightness measurements, can give us information about the shapes of celestial bodies. In the first problem, we discuss an effect where sunlight produces a torque that affects the rotational dynamics of small bodies. The effect was named YORP, after its observers Yarkovsky, O'Keefe, Radzievskii and Paddack. We derive analytical formulas for the YORP torques and a new quantity that expresses the maximum YORP effect caused by the sunlight for a given shape. We called this quantity YORP capacity. In addition, we estimate the upper bound of the YORP capacity, showing how it is theoretically unbounded, but for practical shapes, there is an approximate, finite upper bound. We also study the stability of the YORP effect against noise in the shape model, discovering that the absolute change in YORP capacity remains small with minor shape perturbations, but the relative change becomes large if the YORP capacity was initially small. Analyzing the shape models found in an asteroid database, we discovered that a majority of the shapes is unstable against shape perturbations, while a fraction of the shapes is semistable. The former case may explain why astronomers have obtained completely different results on the YORP effect when using shapes with different resolutions. All of the aforementioned are new results in the field of YORP research.

In the second case, we study an inverse problem where the quality of the data, given as time series, is so weak that it is no longer possible to reconstruct a model for an individual target. Therefore, rather than examining single targets, we focus on a large population of targets, and attempt to obtain information on the population. That is, we attempt to reconstruct a population-wide model. The characteristics of the population are described with a distribution function. This approach is completely new in the field of inverse problems. We discuss the forward model and the inverse problem, showing that even with weak data and a crude model, it is possible to obtain a unique joint distribution that gives us information about our two parameters, shape elongation and spin latitude. The accuracy of the solution is rough, but brings out the information that can be recovered from the data, and the distribution method is tolerant to data and model noise. We introduce some applications of the distribution method, such as a synthetic simulator for estimating the accuracy of the obtained solution, and some ways to measure the differences between the distributions obtained for different populations. In addition, we developed a software package that implements the distribution method and the above-mentioned applications on a user-given asteroid database. The software package can be used for experimenting with different populations, and inspecting different hypotheses or correlations. For example, we confirmed a previous study that the YORP effect has a noticeable effect on the distribution of spin latitudes on a specific asteroid family.

Tiivistelmä

Tässä väitöskirjassa tarkastellaan kahta tutkimusongelmaa. Ensimmäisessä tutkitaan, miten taivaankappaleen muoto vaikuttaa erään valon aiheuttaman ilmiön voimakkuuteen, ja toisessa tarkastellaan, miten valon avulla (kirkkausmittausten kautta) saadaan tietoa taivaankappaleiden muodosta. Ensimmäisessä tutkimusongelmassa auringonvalo aiheuttaa vääntömomentin, joka vaikuttaa merkittävästi erityisesti pienten kappaleiden pyörimisdynamiikkaan. Ilmiö sai nimen YORP havaitsijoidensa Yarkovsky, O'Keefe, Radzievskii ja Paddack mukaan. Väitöskirjassa johdetaan ilmiöön liittyvät analyyttiset kaavat sekä YORP-kapasiteetti: uusi suure, joka ilmaisee suurimman auringonvalon aiheuttaman efektin annettua muotoa kohtaan. Lisäksi väitöskirjassa arvioidaan YORP-kapasiteetin teoreettista ja käytännöllistä ylärajaa ja tutkitaan YORP-ilmiön stabiiliutta tapauksessa, jossa muodon mallinnukseen liittyy kohinaa. Kun eräässä asteroiditietokannassa olevat muotomallit käytiin läpi, havaittiin, että suurin osa muodoista on epästabiileja kohinan suhteen, kun taas noin kolmasosa on semistabiileja. Asteroidien YORP-epästabiilius selittänee, miksi tähtitieteilijät ovat saaneet täysin erilaisia tuloksia mallin tarkkuuden muuttuessa. Kaikki edellä mainitut ovat uusia tuloksia YORP-tutkimuksessa.

Toisessa tutkimuskysymyksessä tarkastellaan inversio-ongelmaa, jossa aikasarjana annetun datan laatu on niin heikkoa, että yksittäisille kappaleille ei pysty enää rakentamaan mallia. Siksi tarkastelu keskittyy yksittäisten kohteiden sijaan suuriin kohdepopulaatioihin, ja malli pyritään yksittäisen kappaleen sijaan muodostamaan koko populaatiolle; tämän populaation ominaisuudet kuvataan tiheysfunktion avulla. Tällainen lähestymistapa on täysin uusi inversio-ongelmien alalla. Väitöskirjassa johdetaan suoran ongelman malli sekä inversio-ongelman ratkaisu ja näytetään, että heikkotasoisesta datasta ja karkeasta mallista huolimatta voidaan yksikäsitteisesti määrittää yhteisjakauma, jonka kautta saadaan tietoa tarkasteltavista parametreistä, pitkulaisuudesta ja pyörimissuunnasta. Ratkaisun tarkkuusluokka on karkea, mutta heikosta datasta saadaan irti kaikki, mikä siitä on palautettavissa, ja lisäksi menetelmä sietää hyvin datan ja mallin kohinaa. Väitöskirjassa esitellään myös distribuutiomenetelmän sovelluksia, kuten synteettinen simulaattori, joka on ainoa keino estimoida saadun ratkaisun tarkkuutta, sekä komponentti, joka vertailee eri populaatioista saatuja ratkaisujakaumia keskenään ja ilmaisee jakaumien eroja. Tutkimusta varten kehitettiin lisäksi ohjelmistopaketti, joka analysoi käyttäjän antamaa tietokantaa distribuutiomenetelmän ja edellä mainittujen sovellusten avulla. Ohjelmistopaketin avulla voi kokeilla erilaisia populaatioita sekä testata erilaisia hypoteeseja tai korrelaatioita. Esimerkiksi väitöskirjassa vahvistetaan aiempi tutkimustulos siitä, että YORP-ilmiö vaikuttaa huomattavasti erään asteroidiperheen pyörimissuuntien jakaumaan.

Preface

The research presented in this thesis was carried out in the Laboratory of Mathematics at Tampere University of Technology (TUT). I started my doctoral studies in the first half of the year 2012, a few months after graduating as a Master of Science at TUT. Professor Mikko Kaasalainen, who would later become my supervisor, presented me with a task of designing an algorithm for generating randomized shapes. Since I had studied mathematical modelling, as well as some scientific computing during my Master's degree, and I was interested in applied mathematics, I found the problem intriguing, and I came up with some creative ideas on how to implement the shape generator. This eventually lead to me joining the research group of the Inverse Problems, and also beginning my doctoral studies.

I would say that the process of doing a dissertation is not linear, or at least it was anything but linear for me. Instead, I would compare it to the process of digging for treasure. For a long time, it may seem like you are not doing any progress, the pit you are digging is only getting deeper, and there seems to be no signs of advancement. Then you suddenly locate the gold, and after that, your motivation will skyrocket, you will dig even harder to reach for the treasure, and the rest of the process will be quite smooth and straightforward. The years 2014 – 2016 were spent experimenting with several topics in the field of inverse problems to come up with a central theme for my doctoral thesis, and there was a long and complicated review process before the first article of the distribution method, Nortunen et al. (2017), was accepted. However, with the acceptance of the paper in the early 2017, the remaining articles were submitted and accepted in a quick pace.

Even if this is a frequently repeated remark, I would like to state it here. When I finished high school, I thought I knew almost everything. When I started my first year at the university, it became obvious to me that this was a misconception. After finishing my Master's degree, I realized that I do not know much at all, if anything. During the process of doing my doctoral thesis, I have studied a field about which most people do not know anything, and some may not understand the field even if I explained it to them. Nevertheless, even I do not think that I know everything of my research topic. Much of it is still beyond my area of expertise or even beyond my understanding. However, the key is not whether you know or understand everything, but what your attitude is towards learning. What I have found important is to keep an open attitude of being willing to learn and try new things, getting to know fields that I had not explored before. I still have a long way to go even towards mastering this attitude, but walking towards this

goal is part of the growth.

> "How in the midst of all this sorrow can so much hope and love endure? I was innocent and certain, now I'm wiser but unsure."
>
> – from *Beauty and the Beast* by Walt Disney Pictures, 2017

I would like to express my most sincere gratitude to my supervisor, Professor Mikko Kaasalainen, for continuous encouragement and valuable comments throughout the doctoral studies. Many of his writing advice have been very helpful as well. Rather than writing everything in a linear order, he suggested that I should first focus on the big picture, and then iteratively proceed towards adding more details. This approach has worked not only with the doctoral thesis, but other writing occasions as well. I am also grateful for his positive attitude of promoting creativity and innovation. In addition, I would like to thank the Academy of Finland (through the Centre of Excellence in Inverse Problems Research, later the Centre of Excellence of Inverse Modelling and Imaging) for their support, as well as Jenny and Antti Wihuri Foundation for their research funding in 2014 – 2016.

I would also like to thank the pre-examiners of my doctoral thesis, Professor Erkki Somersalo and Associate Professor Juha Vierinen, for their valuable comments and feedback, and for their interest towards my research.

I want to thank my colleagues at workplace. Special thanks belong to Markku Åkerblom and Matti Viikinkoski for numerous practical advice with the doctoral thesis. I want to thank Pasi Raumonen, Henri Riihimäki, Sampsa Pursiainen, Mika Takala, Elina Viro and Kalle Rutanen for our moments and discussions together. Last but not the least, I would like to thank Kari Suomela for all the humorous and encouraging conversations we have had during all these years at the Laboratory of Mathematics. For the associates at Charles University in Prague, I want to thank Josef Ďurech, Helena Cibulková and Josef Hanuš for their collaboration, and for generously hosting me on my visits to their university.

As the dissertation process requires plenty of work and time, it is vital to have something to counterbalance all the hard work. I am grateful for the students from Hiukkanen; thank you for having a place where I can wind down after a working day. I want to thank Jenna Mannoja, Martti Bergström and their family for our friendship: Jenna for all the fun and deep conversations, as well as analyzing Marvel and Lord of the Rings stories, and Masi for our funny Tichu games and constantly reminding me how mathematicians cannot count. I would also like to shout "Omstart!" and thank all the friends from NääsPeksi. With all the academic writing and rationality associated with the dissertation, I find it extremely important to keep in touch with my emotional and artistic side as well. The work on our student theater productions has been refreshing and unforgettable, filled with humor and laughter. Among others, I want to thank Sanna Kallio, Pinja Peltohaka, Juha Tiihonen, Santtu Laine, Anna-Maria Vuorinen and Heidi Hansen-Haug for all the

wonderful moments and hardships that we have shared together.

> "Light floods through memories, helps me walk my path, I'll keep my head up high. Words of faith and love, your strength gives me hope."
>
> – from *Final Fantasy XIII-2* by Square Enix, 2011

Haluan kiittää perhettäni heidän tuestaan opintojeni aikana. Äidille ja isälle, Leilalle ja Eerolle, kiitos siitä, että olette jaksaneet uskoa siihen, että pystyn suorittamaan tohtorintutkinnon. Ennen jatko-opintoja mietin pitkään, että diplomityö oli paras, mihin pystyn, eikä minulla osaaminen riittäisi mitenkään tohtorintutkintoon. Kiitos kuitenkin siitä, että te jaksoitte luottaa kykyihini, vaikken itse aina voinut. Kiitos myös siitä, että olette jaksaneet kestää vaikeita, stressaantuneita ja väsyneitä hetkiäni. Haluan myös kiittää sisaruksiani ja heidän perheitään. Kiitos Kaitsulle, että olet hoitanut tietokoneongelmiani ja -asennuksia; sellaiset asiat kuten hajoavat kiintolevyt tai sirkkelin lailla huutava prosessori- tai takatuuletin väsyttävät yllättävän paljon. Kiitos vielä kummitytölleni Sofialle. Prahan työmatkaltahan ei voi palata ennen kuin jotain tuliaisia on löydetty, ja Karvinen sieltä lopulta tarttui mukaan lahjaksi Sofialle.

Finally, I would like to thank the Lord that I have finished up this doctoral thesis.

Tampere, May 2018

Hari Nortunen

Contents

List of symbols and abbreviations

The acronyms used in this thesis are listed below.

YORP	Yarkovsky-O'Keefe-Radzievskii-Paddack
DAMIT	Database of Asteroid Models from Inversion Techniques
WISE	Wide-field Infrared Survey Explorer
Pan-STARRS	Panoramic Survey Telescope and Rapid Response System
LEADER	Latitudes and Elongations of Asteroid Distributions Estimated Rapidly
DF	Density function
CDF	Cumulative distribution function
MATLAB	Matrix Laboratory

The most commonly used symbols in this thesis are listed below.

$\mathcal{Y}$	YORP capacity
τ	YORP torque
p	shape elongation
β	spin latitude
L	brightness
η	brightness variation

List of Publications

I Kaasalainen, M. and Nortunen, H., "Compact YORP formulation and stability analysis," *Astronomy and Astrophysics*, vol. 558, A104, 2013. DOI: 10.1051/0004-6361/201322221.

II Nortunen, H., Kaasalainen, M., Ďurech, J., Cibulková, H., Alí-Lagoa, V., and Hanuš, J., "Shape and spin distributions of asteroid populations from brightness variation estimates and large databases," *Astronomy and Astrophysics*, vol. 601, A139, 2017. DOI: 10.1051/0004-6361/201629850.

III Nortunen, H. and Kaasalainen, M., "LEADER: fast estimates of asteroid shape elongation and spin latitude distributions from scarce photometry," *Astronomy and Astrophysics*, vol. 608, A91, 2017. DOI: 10.1051/0004-6361/201731360.

IV Cibulková, H., Nortunen, H., Ďurech, J., Kaasalainen, M., Vereš, P., Jedicke, R., Wainscoat, R. J., Mommert, M., Trilling, D. E., Schunová-Lilly, E., Magnier, E. A., Waters, C., and Flewelling, H., "Distribution of shape elongations of main belt asteroids derived from Pan-STARRS1 photometry," *Astronomy and Astrophysics*, vol. 611, A86, 2018. DOI: 10.1051/0004-6361/201731554.

Contributions of the authors in the publications:

In I, MK constructed the analytical theory and wrote most of the text. HN assisted in the formulation and writing, created the figures, performed all the computations and wrote the numerical code. Both authors designed the stability tests and defined the YORP capacity and other related concepts, and compiled the conclusions.

In II, HN wrote most of the text, created the figures and the inverse problem procedure, designed and carried out the simulations and data analysis, performed all the computations and wrote the numerical code, and assisted in the analytical formulation. MK constructed the analytical concepts and basis functions and participated in the writing, simulation design, and data analysis. Other authors created the database and assisted in data analysis, and JĎ and HC additionally assisted in the astronomical conclusions.

In III, HN wrote the text, performed the computations, and designed and wrote the software package. MK assisted in drafting the paper and designing the software.

In IV, HC and HN wrote the text, performed the computations, and compiled the conclusions. HC and JĎ designed the databases and data analysis. HN wrote the code and carried out the simulations. MK and JĎ assisted in writing, and designed and tested the concept of period estimation from scarce data. Other authors helped in data analysis and drafting the paper.

1 Introduction

This doctoral thesis discusses first a direct problem, and later an inverse problem, in the field of applied mathematics. Our target of application lies in astronomy, or more accurately, celestial bodies, such as asteroids. Many of the publications related to this dissertation are done as an interdisciplinary, collaborative work with the Astronomical Institute of Charles University in Prague, Czech Republic.

In the direct problem, we discuss the Yarkovsky-O'Keefe-Radzievskii-Paddack (YORP) effect, where sunlight affects the rotation rate and spin direction of small celestial bodies. We focus on the relation between the shape and light, that is, what kind of impact the shape of the body has on the effect of the sunlight. We define a new quantity called YORP capacity. The quantity is dimension-free and depends only on the shape of the body, and it indicates the maximum effect that the light can cause on the body. In our YORP paper (Kaasalainen and Nortunen, 2013), we additionally derive the analytical expressions of the YORP torques, both the component that alters the rotation speed of the body, and one that alters the obliquity; that is, the tilt latitude of the spin direction with respect to the orbital plane. We also estimate the stability properties of the YORP effect against shape perturbations, examining how the YORP capacity changes with perturbations applied on the shape either locally or globally. We noticed that asteroids can be divided into stability classes, where a higher YORP capacity results in a higher YORP stability.

In the inverse problem, we examine a different case study in the field of astronomy. Once again, we focus on the relation between the shape and light, but this time with the following question: how can we obtain shapes of bodies with the help of light, that is, brightness measurements? We present a pathological case where the quality of the brightness data is so weak that it is no longer possible to reconstruct the shape of an individual body. If the situation is so grim, and information on the shape of individual asteroids cannot be recovered, our only hope to obtain information is to consider a population of asteroids, and see if we can recover anything on a population-wide level. We treat the population as a distribution, and we attempt to reconstruct the density function describing the shape distribution of the whole population. Naturally, the targets have lost their identity in the distribution, so even if we can reconstruct the shape distribution of the population, we will be unable to use the distribution to trace back individual targets of the population.

We described our method of using population-wide analysis and distribution functions in Nortunen et al. (2017). In the paper, we constructed the mathematical background of the problem, and formulated an observable, an estimate of brightness variety, which we obtain from the brightness measurements of the population. We proved how the inverse problem has a unique solution, so the joint distribution for the desired characteristics, shape

elongation and spin latitude, can be uniquely determined. We performed simulations on synthetic populations where the elongation and latitude distributions were known, and we discovered that even with our crude model, the joint distribution can be reconstructed on a rough-scale accuracy, and the approach turned out to be robust; that is, even by adding noise in the brightness measurements and trying different models, the observable remained mostly unaffected by the changes.

We tested our method using real astronomical data, using the WISE database in Nortunen et al. (2017), and the Pan-STARRS1 database in (Cibulková et al., 2018). In addition, we created a software package called Latitudes and Elongations of Asteroid Distributions Estimated Rapidly (LEADER), which can compute the joint shape elongation and spin latitude distributions for a given population, perform synthetic simulations to estimate the accuracy and validity of the method for a given asteroid database, and compare distributions of different populations to pinpoint differences between the populations.

1.1 Earlier research

The YORP effect has been under study since the early 1900s, but the research did not reach a breakthrough until late 1990s. The main historical events concerning the YORP effect are explained in Chapter 2. To this date, no complete analytical formulation of the YORP effect had been derived, so in Kaasalainen and Nortunen (2013), we aimed to construct the accurate, analytical theory, and derive the mathematical expressions for the YORP torques. We also wanted to explain the observed instability of YORP computations.

There has been little to no previous research about using population-wide data in order to solve an inverse problem in a setup similar to ours. What makes our setup exceptional is that the data are deficient for creating any kind of model of invididual objects. Furthermore, the observations are indirect in the sense that they do not yield any explicit information on the characteristics we are trying to determine. In the field of astronomy, Szabó and Kiss (2008) and McNeill et al. (2016) proposed an interesting concept where they treated a population as a distribution, and attempted to obtain information about the population this way. Their method seemed promising, but we considered the mathematical foundation for their approach to be insufficient, as they did not include a theoretical analysis of their method, or any validity examination. Inspired by their work, we aimed to expand upon their research, and in Nortunen et al. (2017), we constructed the accurate mathematical groundwork to justify the use of the distribution analysis, as well as testing the validity of the method, and analyzing the stability properties of the inverse problem.

1.2 Objectives of the thesis

In the direct problem, we aim to derive the analytical and numerical methods for estimating the YORP capacity, that is, the maximum effect the light can have on a given shape. Furthermore, we are interested in determining how sensitive the YORP capacity is to shape perturbations. The shape models for objects typically contain modeling errors, so it is important to have an understanding of how much these model deficiencies affect the computed YORP capacity. These research problems are discussed in Kaasalainen and Nortunen (2013) as well as Chapter 2, and the numerical method for the computation of the YORP capacity is presented in Section 2.1. In Section 2.2, we additionally present

how the YORP capacities are distributed in a sample population of asteroids, providing insight on how sensitive typical asteroids are to shape perturbations.

In the inverse problem, our objective is to derive the mathematical theory needed to solve the inverse problem, as well as providing a way to estimate the accuracy of the obtained solution. The accuracy estimation is crucial, since the magnitude of the model and data errors is remarkable, and it is vital to gain an understanding of how accurate the obtained results are. The mathematical principles of the method are explained in detail in Nortunen et al. (2017) and revised in Cibulková et al. (2018), while the computational principles are presented in Nortunen and Kaasalainen (2017). The main results of both the analytical and computational aspects are summarized in Chapter 3. A method for the accuracy estimation is presented in Nortunen et al. (2017), and in greater detail in Nortunen and Kaasalainen (2017), while some additional details are discussed in Chapter 4.

1.3 General background of the methods

Determining the YORP capacity for a given asteroid shape is a **direct problem**. In a direct problem, we start with the causes, and use them to compute the results. In our situation, the given shape is the cause, and the YORP capacity (that is, the maximum YORP effect on the body) is the result. The computation of the YORP capacity is straightforward: with the shape known, we can use explicit formulas to determine the capacity. On the other hand, determining the shape and spin distributions from a set of brightness measurements is an **inverse problem**. In an inverse problem, we start with the results, and use them to compute the causal factors. Inverse problems are typically more complicated than direct problems. In our study, the brightness measurements are our given "results" (observations), and the unknown shape and spin distributions are the cause we are trying to determine.

Some methods for solving various inverse problems are presented in Kaipio and Somersalo (2005). Estimating functions as a sum of basis functions has been discussed by several authors, such as Aster et al. (2012). However, there is little to no prior research about solving a population-wide inverse problem in a situation where information on individual objects is unobtainable, so currently, there is no state of the art method for our problem setup. Eventually, our system of equations becomes a linear least squares problem with a positivity constraint. There are several numerical algorithms available for solving such problems (for example, see Press et al. (2007)), but to mitigate the ill-posedness of the inverse problem, some regularization should be included as well (see Kaipio and Somersalo (2005)).

1.4 Outline of the thesis

This doctoral thesis consists of 6 chapters in total. In this chapter (that is, Chapter 1), we introduce the main topics and themes of this thesis on a general level, focusing more on their relevance to the overall picture of the thesis rather than technical details. We also discuss some of the prior research on the topics.

In Chapter 2, we discuss the YORP effect. We introduce a quantity called YORP capacity, and present an algorithm for the numerical computation of the capacity. The determination of the YORP capacity is a direct problem, and the quantity indicates the maximum capability of the sunlight to alter the rotational dynamics of the surveyed body,

depending on the shape of the body. The YORP capacity can be computed for individual shapes. We also discuss the stability of the YORP effect, and perform a study on how the YORP capacities are distributed on a database containing shape models for numerous asteroids.

In Chapter 3, we discuss a challenging inverse problem, where the quality of the data is insufficient for the reconstruction of characteristics for single objects. By considering object populations rather than single targets, we attempt to reconstruct the population-wide distribution of said characteristics. The data are available as a time series, in the form of brightness measurements. We summarize the theory related to the forward model and the inverse problem, and we also explain the main principles of some of the applications of the distribution method.

Chapter 4 focuses on some of the technical details of the aforementioned distribution method. We study some of the design parameters, such as the choice of grid and the amount of regularization, and use our LEADER software package to estimate the accuracy of the reconstructed distribution with those parameters. This way, we can obtain a good understanding of generally applicable parameter choices, and we also set an example on how a user should initially perform an analysis to estimate the accuracy of the solution with the asteroid database that they are using.

Chapter 5 presents some tests we carried out using data from astronomical databases. For example, we tested whether certain properties of a population caused bias and affected the accuracy of the solution. We also discussed the connection between the YORP effect and the spin distribution of asteroid families, using our distribution method to compare with some of the results obtained by astronomers.

Finally, we summarize and discuss the main results in Chapter 6.

2 YORP: the impact of light, and how it is affected by shape

The **Yarkovsky** effect is a force caused by sunlight that has an impact on the orbital motion of asteroids (Vokrouhlický et al., 2015). While the effect is small, it is significant for small asteroids with a diameter up to 30 − 40 km. The sunlight causes the bodies to heat up, and eventually they radiate the energy away, causing a minor thrust. The phenomenon is effectively as if the asteroids had tiny jets attached to them. While the recoil is minimal compared to the standard solar and gravitational forces, the Yarkovsky effect can cause notable changes in the orbit in a time span ranging from millions to billions of years. The orbital changes become measurable much faster, in a matter of decades.

In addition to the thrust, the sunlight causes a thermal torque. This torque, in addition to a torque created by the scattering sunlight, can produce changes even on the rotation rates and obliquities of small asteroids. This phenomenon of the net torque affecting the rotational behavior of celestial bodies is called Yarkovsky-O'Keefe-Radzievskii-Paddack (**YORP**) effect (Vokrouhlický et al., 2015). Many of the unsolved mysteries in the planetary science of small bodies were solved with the discoveries of Yarkovsky and YORP effects.

The Yarkovsky effect was named after Ivan O. Yarkovsky, a Russian civil engineer, who published a pamphlet in the early 1900s (Yarkovsky, 1901), where he proposed that heating a celestial body can produce an acceleration in the motion of the body. His work was a counter to the ether hypothesis, which was popular at that time. Although Yarkovsky was able to estimate the magnitude of the phenomenon only roughly, his work was expanded later into a complete theory of how solar energy can alter the orbits of small bodies. Yarkovsky's work was kept alive and brought known to the western world in the 1950s by the Estonian astronomer Ernst J. Öpik (Öpik, 1951). Also in the 1950s, Vladimir V. Radzievskii published scientific articles which covered similar topics (Radzievskii, 1952; Radzievskii, 1954), being the first to consider how photon thrusts affect the rotation of a body. Stephen J. Paddack and John O'Keefe expanded upon the concept by discovering that thermal radiation will alter the spin rate of an irregularly shaped celestial body (Paddack, 1969; Paddack and Rhee, 1975). The phenomenon was named the Yarkovsky-O'Keefe-Radzievskii-Paddack (YORP) effect, as a rotational variant of the Yarkovsky effect.

The modern knowledge of Yarkovsky and YORP effects was formed between the 1950s and 1970s. A new interest in the phenomenon was roused in the late 1990s by the works of David P. Rubincam (Rubincam, 1995; Rubincam, 1998, 2000) and Paolo Farinella (Farinella et al., 1998). The research gained a boost, and numerous new results were

obtained, as well as applications to the dynamics of small asteroids and their populations (Bottke et al., 2002, 2006). After that, both Yarkovsky and YORP effects have been under study, and today, both of them are important effects to planetary sciences, as well as being important to the agenda of space missions.

For a technical review of the Yarkovsky and YORP effects, a good overview is presented, for example, in the works of Bottke et al. (2002, 2006). Even though the effect is fairly known and popular these days, no comprehensive analytical theory of the YORP effect had been constructed before. In addition, there had never been a mathematical examination about the theoretical upper bound of the YORP effect. Therefore, in Kaasalainen and Nortunen (2013), we aimed for a thorough analysis of the phenomenon in order to formulate an accurate, analytical theory. We derived the analytical formulas for the thermal YORP torques caused by the emitted photons. There are two main secular components of the YORP torques, one that changes the rotation speed, and one that changes the obliquity. In addition, we carried out numerical experiments to make sure that the results we obtained from them coincided with the analytical theory we formulated.

An interesting question is why the thermal radiation is such a dominant factor in the YORP effect. When a solar photon reaches the body, one of the following three alternatives may occur. The photon may simply scatter off the surface. Alternatively, the photon may be absorbed and cause a torque, or the absorbed photon may be emitted, with the emitted photon causing thermal torque. Since asteroids are mostly black bodies, the ratio of scattering photons is insignificant, with the surface albedo factor A (see Kaasalainen and Nortunen (2013)) being close to 0. Thus, only a small percentage of the light is directly scattered. According to Nesvorný and Vokrouhlický (2008), the torque produced by absorbed photons is given by

$$\tau = -\frac{F}{h^2 v_c} \int_{S^+} (\mathbf{x} \times \mathbf{n}_0)(\mathbf{n} \cdot \mathbf{n}_0) \mathrm{d}S, \tag{2.1}$$

where the integral is taken over the illuminated part of the surface (denoted S^+), F is the solar flux, h is the heliocentric distance of the object, v_c is the speed of light, and $\mathbf{x}$ is the radius vector from the centre of mass. The unit vectors $\mathbf{n}$ and $\mathbf{n}_0$ point from the surface element $\mathrm{d}S$ in normal direction and towards the Sun, respectively. Nesvorný and Vokrouhlický (2008) showed that the torque by the absorbed photons vanishes when averaging over orbit revolution and body rotation. Finally, according to Nesvorný and Vokrouhlický (2007); Nesvorný and Vokrouhlický (2008) and Kaasalainen and Nortunen (2013), the thermal YORP torque produced by emitted photons is given by

$$\tau = -\frac{2\varepsilon_l \sigma}{3 v_c} \int_S (\mathbf{x} \times \mathbf{n}) T^4 \mathrm{d}S, \tag{2.2}$$

where the integral is taken over the entire surface S of the body, ε_l is the emissivity of the material in thermal wavelengths, σ is the Stefan–Boltzmann constant, T is the surface temperature, and v_c, $\mathbf{x}$, $\mathbf{n}$ and $\mathrm{d}S$ are as in (2.1). Averaging (2.2) produces a non-zero torque. Thus, thermal radiation is the only dominating factor, with the other effects produced by the photons being negligible.

It should be noted that for an increasing asteroid size scale factor r, the acceleration slows down in $1/r$ for the Yarkovsky effect, and as a factor of $1/r^2$ for the YORP effect. For the Yarkovsky effect, the acceleration is given by $a = F/m$ according to Newton's second law. The force F is applied on a surface area, which is proportional to r^2, while the mass m is proportional to r^3. Hence, the acceleration has a size dependence of $1/r$. For the YORP

effect, it holds for the torque of a rotating body that $\tau = I\alpha$, or $\alpha = \tau/I$, where α is the angular acceleration, and I is the moment of inertia. The torque τ is proportional to the product of the force applied to the lever, and the length of the lever arm. The force, as above, is proportional to r^2, and with the length of the lever arm being proportional to r, the torque τ is proportional to r^3. Meanwhile, it holds for the moment of inertia that $I \propto mr^2$, so with the mass being proportional to r^3, I is proportional to r^5. With $\alpha = \tau/I$, it follows that the angular acceleration has a size dependence of $1/r^2$. This means that if body A is n times bigger than body B, then the YORP effect of body A is n^2 times weaker. This explains why the YORP effect is typically observed for only small bodies.

Our focus, however, is not the size dependency of the YORP effect, but the shape dependency. In order to determine how much the shape of the body affects the YORP torques, we defined a new quantity, YORP capacity $\mathcal{Y}$. To obtain $\mathcal{Y}$, we scale the YORP torques with the volume of the body and other physical quantities to make them size-invariant and dimensionless. For completely symmetric bodies, the YORP torques vanish, and thus, $\mathcal{Y}_{\min} = 0$. To estimate the scale of the YORP capacity, we considered the shape of a propeller. Assume we have two identical wedges that are vertically symmetric and attached to each other (see Kaasalainen and Nortunen (2013)). Then, for an infinitely long and thin body, we get $\mathcal{Y}_{\max} \to \infty$, so there is no theoretical upper bound for the YORP capacity. However, for realistic shapes, we can estimate that $\mathcal{Y} \lesssim 0.1$.

Another quality we are interested in is the YORP stability of shapes. If the shape is only slightly different for two bodies, how much difference do the distinctions in shape cause in the YORP properties of the bodies? In many asteroid models, such as the ones in the Database of Asteroid Models from Inversion Techniques (**DAMIT**[1], the bodies are approximated as convex shapes. The validity of this approximation depends on how stable the YORP capacity is to perturbations in shape. Furthermore, if we have multiple models of the same body, how much will the differences in the models affect the YORP capacity, especially if the models have different resolutions? In Kaasalainen and Nortunen (2013), we estimated the sensitivity of the YORP capacity to shape perturbations both analytically and numerically. Let $r(\theta, \phi)$ be the radius of the body given in spherical coordinates. If we apply a global perturbation on a shape, for example, by using a low-order Laplace series, we showed how the change in YORP capacity, $\Delta\mathcal{Y}$, has an upper bound:

$$|\Delta\mathcal{Y}| \lesssim k_{\max} \frac{\langle \Delta r \rangle}{\langle r \rangle}, \tag{2.3}$$

where $k_{\max}$ is a constant and Δr is the change in radius. This means the absolute change in YORP capacity will remain bounded, despite the perturbations. For most shapes, $k_{\max} \sim 0.1$. Let us now consider the relative change. Assuming the relative model error $\langle \Delta r \rangle / \langle r \rangle$ remains small, shapes with $\mathcal{Y} \gtrsim 0.01$ can be estimated to be reasonably YORP-stable to global model errors, with the relative YORP error $|\Delta\mathcal{Y}/\mathcal{Y}|$ remaining small. That is, any instabilities in the YORP capacity are caused by the $\mathcal{Y}$ value being initially small, rather than the size of $|\Delta\mathcal{Y}|$.

In additional to the global perturbations by a Laplace series, we studied the effect of local perturbations, implemented by applying a simple perturbation on one of the vertices of the model. The local perturbations represent a local model error, such as a boulder or a crater, or an actual local change on the shape, caused by a physical perturbation,

[1] http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php (cited on April 30th, 2018)

such as a collision or erosion. We noticed that shapes tend to be more stable to local perturbations, even if they are not globally stable to perturbations. For shapes with $\mathcal{Y} \lesssim 0.001$, however, the shapes are unstable even against local shape perturbations. In Sect. 2.2, we discuss the YORP stability of actual asteroids, based on their YORP capacity.

The analytical theory with the YORP torque formulas has been presented in Kaasalainen and Nortunen (2013), while the algorithm for the numerical computation of the YORP capacity is covered in Sect. 2.1. We computed $\mathcal{Y}$ for each asteroid model found in the DAMIT database to obtain an understanding of the scale of $\mathcal{Y}$ for realistic shapes; these results are presented in Sect. 2.2.

2.1 Technical implementation for the numerical computation of $\mathcal{Y}$ values

We compute the numerical value for the YORP capacity $\mathcal{Y}$ using the algorithm below. The data files containing the shape models of asteroids were obtained from DAMIT.

1. We read the list of vertices and faces from the data object file, and use the vertex and face information to plot the shape of the asteroid model.

2. We compute the coefficients f_l^m and g_l^m of the spherical harmonics expansion of the radius numerically (see Appendix A.1).

3. We may optionally add perturbations (that is, noise) to the shape, either by applying a pointwise perturbation on some of the vertices (local perturbations), or by perturbing the coefficients of the spherical harmonics expansion (global perturbations). The former is a simpler method and useful for small, controlled perturbations, but it has the drawback of possibly creating unrealistic sharp edges on the surface. The latter creates a smoother body, but the coefficients of the spherical harmonics expansion are very sensitive to perturbations, so a tiny change in the coefficients may completely alter the shape of the body.

4. For each face, we compute the unit normal, area, volume, center point and center of mass. With the center of mass computed for each face, we fix the total center of mass of the complete shape into the origin.

5. All direction vectors v must be transformed from the orbital frame to the asteroid-fixed frame:
$$v_{\text{ast}} = R_z(\varphi')R_y(\epsilon)v_{\text{orb}}, \tag{2.4}$$
where φ' is the rotation angle of the body around its axis, ϵ is the obliquity of the rotation axis measured from the orbital pole, and R_y and R_z are rotation matrices for y- and z-axes (Kaasalainen and Lamberg, 2006).

6. Let ω be the direction of the illumination source (in the asteroid-fixed frame). For nonconvex bodies, we need ray tracing to evaluate which faces are illuminated. For convex bodies, the visibility check is much simpler, as the face i is illuminated if $\mu_i := n_i \cdot \omega > 0$, where n_i is the unit normal of face i.

7. With the position of the asteroid around the Sun given by $\lambda \in [0, 2\pi)$, the obliquity of the rotation axis from the orbital pole given by $\epsilon \in [0, \pi]$, and the rotation angle

of the asteroid around its axis given by $\varphi' \in [0, 2\pi)$, the YORP torque can be computed as follows:

$$\tau(\lambda, \epsilon, \varphi') = \sum_{\substack{i;\ \text{face } i \\ \text{is visible}}} A_i \mu_i (x_i \times n_i), \tag{2.5}$$

where A_i is the area and x_i is the centroid of face i. We average over φ' and λ in order to obtain $\tau(\epsilon)$. Finally, we define the YORP capacity $\mathcal{Y}$ as the amplitude of the z-component of $\tau(\epsilon)$ scaled by the physical constant Q and the volume V:

$$\mathcal{Y} := \frac{\max |\tau_z(\epsilon)|}{QV}. \tag{2.6}$$

We are only interested in the z-component of τ because the z-component of the torque causes the biggest changes in the rotation rate of the body, as shown in Kaasalainen and Nortunen (2013). For the definition of $\mathcal{Y}$, we chose the maximum of τ_z instead of the average since we wanted to compute the maximum effect caused by the sunlight on the body. Since we do not take the size and mass of the asteroid into account, we simply set $Q = 1$, and consider only the shape of the asteroid.

In Kaasalainen and Nortunen (2013), we compared the numerically computed $\mathcal{Y}$ values with the analytical $\mathcal{Y}$ for several asteroids to ensure that our results were accurate.

2.2 YORP experiments with asteroid models

With the numerical algorithm described in Sect. 2.1, we can compute YORP capacities for given asteroid models. The DAMIT database contains hundreds of convex asteroid models, with a rich variety of realistic shapes, so the data models give us a good understanding of what kind of YORP capacities are typical for common asteroid shapes.

A thorough study of the YORP capacities for the shapes given in DAMIT was carried out by Miinalainen (2014). The inspection was done not only to obtain insight on the distribution of YORP capacities for different shapes, but also to examine possible correlations between YORP capacity and the size of the asteroid. We note that the YORP capacity is intrinsically size-invariant and only considers the shape of the body, not the size. A 3D histogram showing the distribution of asteroid shapes from DAMIT for different YORP capacities and diameters has been plotted in Fig. 2.1 (from the study by Miinalainen (2014), reproduced with permission).

Clearly, a vast majority of the asteroid models have a small diameter and a small YORP capacity. What is significant is that the highest YORP capacities are encountered with small bodies, while large bodies only have low $\mathcal{Y}$ values. This is realistic, since it is more likely that a small body has an irregular and asymmetric shape, whereas large bodies tend to be more spherical and symmetric. While the YORP capacity itself depends only on the shape of the body and not the size, there is an obvious correlation between the shape and the size.

Among DAMIT asteroid models, the highest YORP capacity was encountered with the asteroid **Alisondoane** (asteroid number 7517), with $\mathcal{Y} = 44.02 \cdot 10^{-3}$, while **Freia** (asteroid number 76) had the lowest YORP capacity, $\mathcal{Y} = 0.2347 \cdot 10^{-3}$. We have plotted both asteroid models in Fig. 2.2. The high $\mathcal{Y}$ value of Alisondoane is caused by the

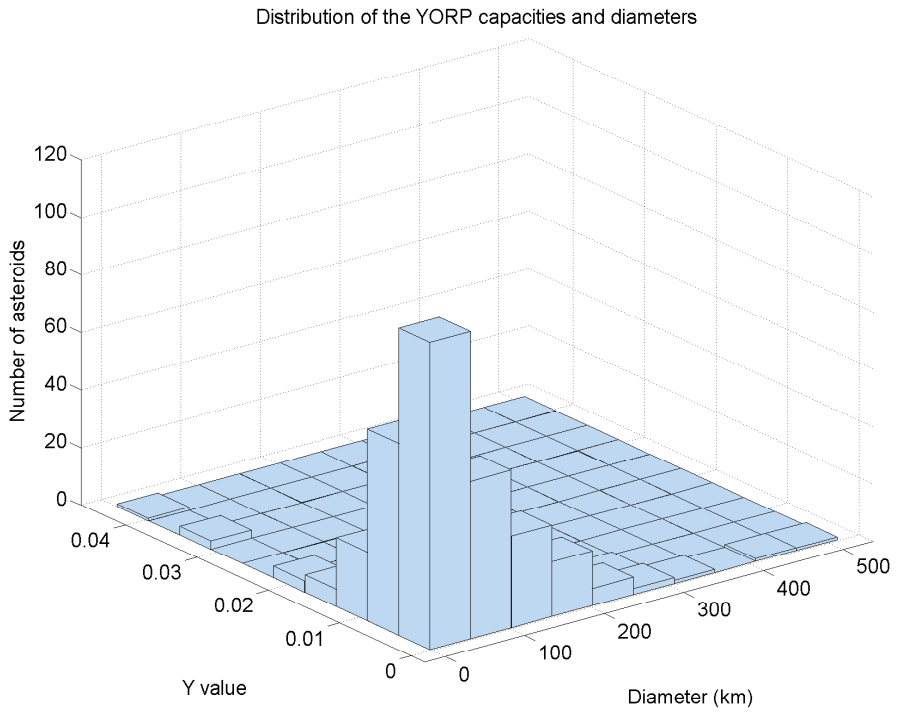

Figure 2.1: Histogram showing the number of asteroids with different $\mathcal{Y}$ values and diameters.

elongated and asymmetric shape of the asteroid, whereas the spherical and symmetric shape reduces the $\mathcal{Y}$ value of Freia considerably, as the symmetries cause the YORP torques to vanish.

Figure 2.2: Asteroid models from DAMIT, with Alisondoane (left) having the highest YORP capacity, and Freia (right) having the lowest YORP capacity.

We mentioned earlier in Chapter 2 that the stability of the YORP effect depends only on the magnitude of the YORP capacity $\mathcal{Y}$. For small perturbations in shape, the absolute change in YORP capacity, $|\Delta\mathcal{Y}|$, remains bounded, while the relative change $|\Delta\mathcal{Y}/\mathcal{Y}|$

becomes remarkably large for shapes with a low $\mathcal{Y}$ value. Now that we have obtained a distribution of YORP capacities for a sample of shapes, we can divide asteroid shapes into different YORP stability classes according to their $\mathcal{Y}$ values. In Kaasalainen and Nortunen (2013), we examined the stability of the YORP effect for shapes with different $\mathcal{Y}$ values, applying either global or local shape perturbations, to estimate the absolute and relative errors in $\mathcal{Y}$ against perturbations.

Theoretically, propeller-type shapes ($\mathcal{Y} \gtrsim 0.05$) are both globally and locally YORP-stable, but naturally, these shapes are rarely encountered. The shapes with high $\mathcal{Y}$ values in Fig. 2.1 ($0.005 \lesssim \mathcal{Y} \lesssim 0.05$) are globally semistable and locally stable. About 32% of the sample belong to this stability class. Some examples of the stability class include asteroids like Apollo ($\mathcal{Y} = 6.61 \cdot 10^{-3}$), Eger ($\mathcal{Y} = 8.27 \cdot 10^{-3}$), Geographos ($\mathcal{Y} = 5.98 \cdot 10^{-3}$) and Massalia ($\mathcal{Y} = 5.95 \cdot 10^{-3}$), with their $\mathcal{Y}$ values computed in Kaasalainen and Nortunen (2013). Most asteroids (about 66% of the sample) are globally YORP-unstable and locally semistable.

Finally, asteroids with $\mathcal{Y} \lesssim 0.0005$ (about 2% of the sample) are both globally and locally YORP-unstable. For shapes with $\mathcal{Y} \lesssim 0.0001$, all YORP estimates are likely to be invalid due to major instabilities. In cases like these, only upper bounds can be given. For example, the YORP effect of **Itokawa** (asteroid number 25143) was studied by Scheeres et al. (2007), Ďurech et al. (2008) and Breiter et al. (2009), but there were considerable fluctuations in the observed YORP torques depending on the model used, with higher-resolution models completely altering the results. Hence, Itokawa can be considered an example of a highly YORP-unstable shape. In such cases, estimates of YORP torques will be chaotic, and only the approximate magnitude can be obtained.

2.3 Conclusions

In this chapter, we have summarized the history of Yarkovsky and YORP effects. Briefly, the Yarkovsky effect is a thermal propulsion force that alters the orbital motion of celestial bodies, while the YORP effect is a thermal torque that affects the rotational dynamics of celestial bodies. Both effects are caused by the sunlight, and the effect is significant for small asteroids over long time spans. We introduced a quantity called YORP capacity, which we defined in Kaasalainen and Nortunen (2013). The YORP capacity $\mathcal{Y}$ is scaled to be dimension-free and size-invariant, and it only depends on the shape of the body. In Sect. 2.1, we presented a numerical algorithm for the computation of $\mathcal{Y}$.

In addition, we have discussed the stability of the YORP effect. For small perturbations of shape, the absolute change in $\mathcal{Y}$ remains bounded, but if the $\mathcal{Y}$ value is intrinsically small, the relative change in $\mathcal{Y}$ becomes large. Therefore, bodies with a small YORP capacity are unstable to shape perturbations. In Sect. 2.2, we represented a histogram showing the distribution of YORP capacities and diameters for asteroid models from the DAMIT database. Since small bodies tend to be more asymmetric and irregularly shaped, while large bodies tend to be more symmetric and spherical, the small bodies are more likely to have a high YORP capacity. With the $\mathcal{Y}$ values known for different asteroid models, we can divide asteroids into different YORP stability classes. Most asteroids are YORP-unstable to global shape perturbations, while a bit less than one third of the sample turned out to be semistable to global shape perturbations. A low YORP capacity for an asteroid may explain why the estimates of YORP strength can change completely when using asteroid models with different resolutions.

3 Shape and spin distributions for large populations

One of the key topics of this doctoral thesis is the study of distribution analysis in order to reconstruct the solution of an inverse problem where the available data are extremely limited. We have a few points of data as a time series, but it is not enough for the reconstruction of a single, individual target. Instead, we focus on a population-level analysis, where each target has a few data points available. If the population is large enough, we can use analytical basis functions that yield the distributions of the characteristics of the population. While information on individual targets is lost and cannot be recovered, at least the characteristics of the population can be reconstructed. The idea of using distributions rather than observations of a single target to solve an inverse problem is a new approach. There is very little (if any) prior research on the methodology.

We discussed the method of a population analysis in three different papers, with a case study on reconstructing the joint distribution of shape elongation and spin latitude for large asteroid populations. In Nortunen et al. (2017), we explained the theoretical background of the method. The available data are insufficient for the reconstruction of individual asteroid shapes. Szabó and Kiss (2008) proposed an innovative idea of a population-wise analysis in order to obtain information on the distribution of shapes. However, their approach did not cover the theoretical aspects of the method, or any validity tests and analysis (in fact, our results do not agree with theirs). Therefore, we wanted to expand upon their research, and we aimed to formulate a detailed theory on the reconstruction of a distribution of characteristics from very limited data. In addition, Szabó and Kiss (2008) simply assumed the spin latitude to be isotropic, while we aimed to solve the distribution of the spin latitude.

In Nortunen and Kaasalainen (2017), we introduced the **LEADER** software package, which is an implementation of the method we presented in Nortunen et al. (2017). The main routine is designed for reconstructing the joint (shape, spin) distribution of an asteroid population. There is a routine for synthetic simulations in order to test whether the method is accurate and stable with a user-defined database. The differences between databases are noticeable, and some databases seem to provide more accurate results than others. For some asteroid databases, such as **WISE** (Wide-field Infrared Survey Explorer), the method is robust and provides accurate results on a rough-scale level, and is therefore useful for detecting distribution peaks for subpopulations the data of which are given in the WISE database. Finally, there is a routine for comparing the shape elongation and spin latitude distributions of two populations; this routine was especially used in Nortunen et al. (2017) and Cibulková et al. (2018).

Finally in Cibulková et al. (2018), we applied the LEADER software package on the **Pan-STARRS1** (Panoramic Survey Telescope and Rapid Response System) database. First, we performed synthetic simulations in order to test the reliability of our method, which should always be done when using a new database; no database should be blindly used with the software package. We observed that the shape elongation is obtained accurately, while the spin latitude suffers from stability problems. After evaluating the validity of our method, we compared shape elongation distributions of several asteroid families using data from the Pan-STARRS1 database.

With **sparse data**, we refer to observations that have been measured in intervals where the sampling frequency may be much lower than the characteristic frequency of the system. For a time span of one period, the data do not provide us any information, but if we have data available from a long time scale, resulting in numerous observations in various geometries, it can be proved (Kaasalainen (2004), Ďurech et al. (2009), Ďurech et al. (2016)) that we can obtain a unique model that yields the shape of the body, and the rotation period of the target. Since the rotation direction remains the same, and the viewing geometries are known, these factors actually set very strict boundary conditions, and as a result, there is only one period that satisfies the boundary conditions. When the period has been determined, the shape can be obtained via fitting.

For our observations, the situation is much weaker: the observation angles are placed so irregularly that it is not possible to reconstruct the shape for individual targets. We refer to such situations as **scarce data**, to distinguish them from sparse data. An illustrative example of dense, sparse and scarce data sets has been presented in Fig. 3.1. With dense data, we observe a sinusoidal function. With sparse data, we cannot observe the sinusoidal behavior directly, due to the low sampling frequency, but with the large number of observations in various geometries (we note the longer time scale in the top right plot), we can obtain a unique fit. In our example with scarce data, we have several measurements from a certain geometry during a short time span, but this ultimately gives us no information on individual objects since the number of geometries is too low. If the data are so scarce that the properties of individual targets cannot be reconstructed, an interesting research question is if there is any information that can be recovered.

This chapter discusses the distribution method where the spin latitude and shape elongation distributions for large asteroid populations are reconstructed from scarce data. We summarize the theoretical background and the software implementation, which were discussed in Nortunen et al. (2017) and Nortunen and Kaasalainen (2017), respectively. In addition, we cover some details that we examined during the process of writing the aforementioned papers, but which we did not include in the actual publications.

3.1 Main algorithm and its theoretical background

The general method of the LEADER software package is based on a new type of an inverse problem. While inverse problems typically consist of using data to reconstruct something like an image, here our number of data points is so low that we do not have enough information for a proper reconstruction of an object. In a traditional inverse problem, we can improve the "image" by obtaining additional measurements from, say, various angles. We do something analogical in our method: we take data points measured from a large number of objects, such as a population. Then, instead of attempting to reconstruct the properties of an individual object, we aim to reconstruct the statistical properties of the entire population. We treat the properties of the entire population as

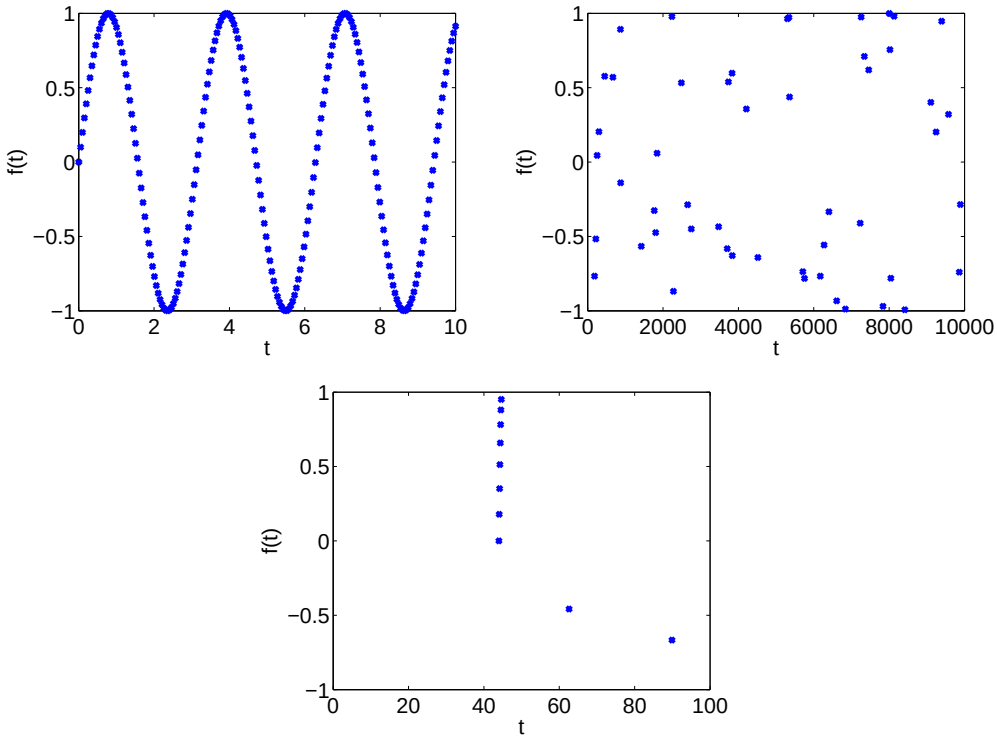

Figure 3.1: Schematic examples of dense (top left), sparse (top right) and scarce (bottom) time series. Note the different time scales on each plot.

a distribution function, and our objective is to reconstruct the DF. If the population is large enough, then we will be able to reconstruct the DF, and so, we will be able to model the population.

Individual objects are lost in this process, that is, we will not be able to identify individual objects from the solved distribution. Similarly to how our typical solution for single targets is an "image", here our solution is an "image" of a population. After obtaining a solution, it is common to apply a posterior correction to enhance or smoothen the picture. Similarly in our method, a deconvolution filter can be applied on the population-scale "image" in order to correct the obtained picture. Our method is not intended for obtaining high-resolution "images" of the population, but it is sufficient in a rough-scale level. The method is also robust, and computationally fast.

3.1.1 Forward model: brightness variation estimate

We attempt to reconstruct two parameters: shape elongation p, and spin latitude β (measured in a coordinate system where the xy-plane is the plane of the Earth's orbit). We are only limited to scarce data, so we do not have enough data per target to obtain the shape of individual objects; a lightcurve inversion for individual bodies would require hundreds of data points for a target. Therefore, rather than focusing on individual targets, as in Chapter 2, we consider an entire asteroid population, and attempt to reconstruct the shape and spin distributions, $f(p)$ and $f(\beta)$, of the population. The reconstruction

typically requires a population of at least a thousand targets in order to obtain an accurate solution.

For the shape elongation, it holds that $0 < p \leq 1$, where $p = 1$ presents a completely spherical body, whereas a p-value closer to 0 presents a "thin cigar". As for the spin latitude, $0 \leq \beta \leq \pi/2$, where $\beta = 0$ presents a direction perpendicular to the ecliptic plane, and $\beta = \pi/2$ presents a direction in the ecliptic plane. We note that for mathematical purposes, β is measured from the pole. This is different from the convention typically used in physics, where the complementary angle of β is traditionally used. In addition, we note that our model cannot distinguish whether the spin latitude is above or below the ecliptic plane.

We approximate the shape of an asteroid with an ellipsoid model, with semiaxes $a \geq b = c = 1$, and we define the shape elongation as the ratio of the two biggest semiaxes, $p = b/a$. Illustrative examples of ellipsoids with different b/a values have been plotted in the xy-plane in Fig. 3.2. We acknowledge that the model is extremely crude, but we use it due to its simplicity[1].

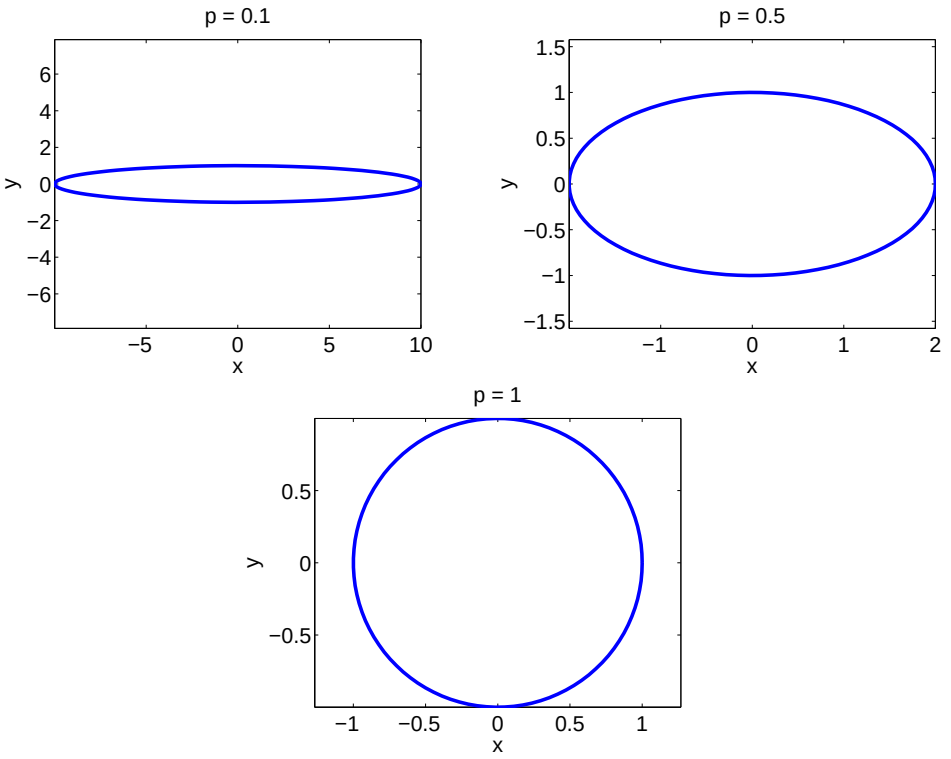

Figure 3.2: Schematic examples of different $p = b/a$ values for the ellipsoid model in the xy-plane, with $p = 0.1$ (top left), $p = 0.5$ (top right) and $p = 1$ (bottom).

Let θ be the polar aspect angle of the viewer, $0 \leq \theta \leq \pi/2$ (it is sufficient to consider this interval instead of $[0, \pi]$ due to model symmetry), and let ϕ be the longitudinal angle in

[1]Other choices of model are discussed in Nortunen and Kaasalainen (2017) and in Sect. 3.3, but we discovered that they did not improve the accuracy of the solution.

the ellipsoid-fixed coordinate frame. The brightness L is given by

$$L = \sqrt{1 + (p^2 - 1)\sin^2\theta\cos^2\phi}. \tag{3.1}$$

for the case of coinciding viewing and illumination directions (that is, the solar phase angle is zero). We may apply phase correction on the observed lightcurves to compensate for the systematic brightness drop if the phase angle increases notably within a sample. Let α be the phase angle between the Sun and the Earth. If the phase angles of the observations are small (we typically used a condition that the phase angles α must be smaller than $8°$) and change for more than one degree in a sample, we correct the observations with a linear curve. For larger angles ($\alpha \geq 8°$), we use an exponential curve for correction if the phase angle changes for more than two degrees in a sample. A more detailed explanation is given in Kaasalainen et al. (2001), and Nortunen and Kaasalainen (2017). The fitted correction curve may be unstable if there are not enough data points in a sample. Therefore, we typically omit the correction if we have less than, for example, five data points. In addition, our model requires that the phase angle α is close enough to zero degrees for the whole sample. Fortunately, the error caused by non-zero phase angles is negligible compared to other error sources. Therefore, we are very liberal with this requirement and only omit measurements if the phase angle α is larger than, for example, 20 or 30 degrees.

Since we want to avoid large changes in geometries, we divide the brightness measurements of a single object into multiple sets, $i = 1, \ldots, S$. For set i, we take all the measurements within a time span $[t_i, t_i + T]$, where T is a fixed time. For WISE database, we noticed that the phase angle never changes for more than $0.4°$ within brightness measurements for a single asteroid, so we are free to choose a large T, such as a month or two. For Pan-STARRS1 database, however, the phase angle changes rapidly in time. The choice of T is arbitrary, but we aim to ensure that the phase angle does not change much within the set, while the set should still contain enough brightness measurements. For Pan-STARRS1, we observed that choosing T to be three days is usually a balanced choice. Nevertheless, we emphasize that the choice of T is arbitrary and depends on the database.

Our main observable is the brightness variation estimate η. Using data from lightcurve brightness measurements, we define η as

$$\eta = \frac{\Delta(L^2)}{\langle L^2\rangle}, \tag{3.2}$$

where L is our observed brightness values, $\langle L^2\rangle$ is the average of the squared brightness L^2:

$$\langle L^2\rangle = \frac{1}{2\pi}\int_0^{2\pi}\Big(1 + (p^2 - 1)\sin^2\theta\cos^2\phi\Big)\,\mathrm{d}\phi = 1 + \frac{1}{2}\sin^2\theta(p^2 - 1), \tag{3.3}$$

and $\Delta(L^2)$ is a measure of variation for L^2, which we defined in Nortunen et al. (2017):

$$\Delta(L^2) = \sqrt{\langle (L^2 - \langle L^2\rangle)^2\rangle} = \frac{1}{\sqrt{8}}\sin^2\theta(1 - p^2). \tag{3.4}$$

We have used the squared brightness L^2 instead of simply L for mathematical convenience; a more detailed explanation is given in Nortunen et al. (2017). We require at least five brightness measurements in a set for a valid η estimate. If we do not obtain enough measurements within a set, we may combine multiple sets into one, as long as we scale the

brightness values of each combined set on the same level (for example, by using average values of each set). As a word of caution, however, we note that we should never combine sets where the geometries are noticeably different; for example, if there has been several months between the sets.

From (3.2), we get that

$$\eta = \frac{1}{\sqrt{8}} \left[\frac{1}{\sin^2 \theta (1 - p^2)} - \frac{1}{2} \right]^{-1}. \tag{3.5}$$

We define the amplitude of the brightness curve as the ratio $A = L_{\min}/L_{\max} = L_{\phi=0}/L_{\phi=\pi/2}$. Thus, from (3.1), we obtain that

$$A = \sqrt{1 + (p^2 - 1) \sin^2 \theta}. \tag{3.6}$$

We note that the amplitude is based on intensity; we do not use magnitudes, which is a typical convention in physics.

In Nortunen et al. (2017), we derived the relation between the brightness variation η and the amplitude A:

$$\eta = \frac{1}{\sqrt{8}} \Big(\frac{1}{1 - A^2} - \frac{1}{2} \Big)^{-1} \iff A = \sqrt{1 - \Big(\frac{1}{\sqrt{8}\eta} + \frac{1}{2} \Big)^{-1}}. \tag{3.7}$$

It should typically hold that $0 \leq \eta \leq 1/\sqrt{2}$. However, when p is small (especially in cases where $p \ll 0.5$), it sometimes occurs that $\eta > 1/\sqrt{2}$, which results in a complex-valued amplitude in Eq. (3.7). In realistic situations, it usually holds that $p \gtrsim 0.5$, so it is rare to obtain complex amplitudes. As the complex-valued are purely a computational concept, we omit them from the set of amplitudes. The loss in data amount caused by complex amplitudes should be insignificant.

When we have collected all of the amplitudes obtained from the brightness measurements of the population, we can construct the cumulative distribution function (CDF) of A, $C(A)$.

3.1.2 Analytical basis functions

In Nortunen et al. (2017), we derived how we can express the CDF $C(A)$ in the form of an analytical integral, assuming we have infinite observations available in every geometry. Let us now divide the (p, β) grid of the population into $n = lm$ distinct bins, (p_i, β_j), where $i = 1, \ldots, l$ and $j = 1, \ldots, m$. In order to determine the distribution for the spin latitude β, it is required that the observations are concentrated near the ecliptic plane. Now, with our discrete (p_i, β_j) grid, we can write the CDF as

$$C(A) = \sum_{ij} w_{ij} \, F_{ij}(A), \tag{3.8}$$

where $F_{ij}(A)$ are monotonously increasing basis functions $F_{ij}(A)$ with the range $[0, \pi/2]$:

$$F_{ij}(A) = \begin{cases} 0, & A \leq p_i \\ \frac{\pi}{2} - \arccos \frac{\sqrt{A^2 - p_i^2}}{\sin \beta_j \sqrt{1 - p_i^2}}, & p_i < A < \mathcal{F}(p_i, \beta_j) \\ \frac{\pi}{2}, & A \geq \mathcal{F}(p_i, \beta_j), \end{cases} \tag{3.9}$$

where $\mathcal{F}(p_i, \beta_j) = \sqrt{\sin^2 \beta_j + p_i^2 \cos^2 \beta_j}$. The coefficients w_{ij} are the occupation numbers w_{ij} of each (p_i, β_j) bin. We note that a case where the occupation numbers are proportional to $\sin \beta$ in the β-direction presents a situation where the spin latitudes are uniformly distributed on a sphere; that is, $f(\beta)$ is constant. It is intuitive to hypothesize that β is uniformly distributed, and for populations with large asteroids, the spin latitudes tend to be more isotropic. However, for populations with smaller asteroids, the YORP effect tends to shift the β peak away from the ecliptic plane, towards the pole.

In realistic scenarios, we do not have an infinite number of observations available, and with the scarce data, the observations are available in a limited set of geometries. For a typical asteroid, we can obtain 1–2 brightness variation estimates, with roughly 5–10 data points for each estimate. Clearly, one or two η values will not yield a smooth CDF. However, encouraged by the law of large numbers, we decided to examine whether the CDF becomes smooth when we take all the η estimates from a large population. Indeed, if we have a large number of asteroids in the population and we can obtain η estimates from each of them, the observed CDF will converge towards the theoretical CDF. For a smooth $C(A)$, we need to compute η estimates for at least a 1000 asteroids, preferably more; see Fig. 3.3 for a demonstration on how the smoothness of the CDF improves as the number of asteroids increases. In other words, we require observations from at least a thousand targets in order for the forward model to become accurate.

3.1.3 Inverse problem: the joint distribution for shape elongations and spin latitudes

The goal of the inverse problem is to obtain the joint distribution $f(p, \beta)$ for the shape elongation $p \in [0, 1]$ and spin latitude $\beta \in [0, \pi/2]$. As in Sect. 3.1.2, we use a grid of $n = lm$ points for (p_i, β_j), $i = 1, \ldots, l$ and $j = 1, \ldots, m$. We can express the superposition of Eq. (3.8) in an equivalent linear form:

$$Mw = C, \tag{3.10}$$

where

$$M = \begin{bmatrix} F_{11}(A) & \ldots & F_{1m}(A) & \ldots & F_{l1}(A) & \ldots & F_{lm}(A) \end{bmatrix}, \tag{3.11}$$

the column vector C contains the values of $C(A)$, and the solution vector w holds the occupation numbers w_{ij} that we attempt to determine. In Nortunen et al. (2017), we proved that the basis functions F_{ij} of (3.8) are linearly independent. Therefore, there is a unique solution to the linear system of (3.10), and since the occupation numbers are uniquely obtained, the joint distribution $f(p, \beta)$ can also be unambiguously reconstructed.

When solving the system (3.10), it is strongly advisable to use regularization, especially for β, which is the more unstable of the two parameters. Let R_p be an $(l-1)m \times n$ matrix intended for smoothing the solution in the p-direction, and R_β the respective $l \cdot (m-1) \times n$ regularization matrix in the β-direction. For indices i and j, we use the following regularization:

$$(R_p)_{ij} = \begin{cases} -1/(p_{i+1} - p_i), & i = j \\ 1/(p_{i+1} - p_i), & j = i + 1 \\ 0, & \text{elsewhere} \end{cases} \tag{3.12}$$

and similarly for R_β. The regularization matrices approximate the gradients in p- and β-directions at each occupation level w_{ij}, attempting to keep the gradients small, and

therefore reducing the changes between adjacent bins; that is, making the solution smoother.

To obtain the occupation numbers w_{ij}, we compute the solution to the following optimization problem:

$$\hat{w} = \arg\min_{w} \left(\|C - Mw\|^2 + \delta_p \|R_p w\|^2 + \delta_\beta \|R_\beta w\|^2 \right), \quad w \in \mathbb{R}^n_+. \tag{3.13}$$

In order to obtain $\hat{w}$, we create an extended matrix $\tilde{M}$ and an extended vector $\tilde{C}$:

$$\tilde{M} = \begin{bmatrix} M \\ \sqrt{\delta_p} R_p \\ \sqrt{\delta_\beta} R_\beta \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C \\ 0_{(l-1)m} \\ 0_{l(m-1)} \end{bmatrix}, \tag{3.14}$$

Now, our extended system is of the form

$$\tilde{M} w = \tilde{C}. \tag{3.15}$$

Due to the instability of the problem, the traditional least squares solution would often lead to some occupation numbers w_{ij} becoming negative, and it is also common that some occupation numbers blow up, approaching $\pm\infty$. Non-linear optimization with a change of variables, for example $w_{ij} = \exp(z_{ij})$, would ensure the positivity of the occupation numbers, but we found the method too inaccurate and computationally slow. Instead, we used MATLAB's linear least squares method with a positivity constraint $w_{ij} \geq 0$ (Lawson and Hanson, 1974); this method turned out to be effective. Now, let $\tilde{w}_{ij}$ be the occupation numbers w_{ij} normalized such that $\sum_{i,j} \tilde{w}_{ij} = 1$. To obtain the joint distribution for p and β, we can simply write $f(p_i, \beta_j) = \tilde{w}_{ij}$.

We tested different population sizes to see how large population is required for the function series $\sum_{ij} w_{ij} F_{ij}$ to converge towards the CDF $C(A)$. In Fig. 3.3, we see that a data sample of 10 asteroids is obviously insufficient for convergence. With a sample of 100 asteroids, the shape of the CDF is beginning to show, but the CDF is still poorly sampled. With 1000 asteroids, the CDF is smooth, and the function series fitting is accurate, with an error of less than 1%.

We assume that the observations are concentrated near the ecliptic plane (otherwise there is no information on the spin latitude). We tested the validity of the assumption, and for WISE and Pan-STARRS1 databases, about 95% of the observations were located in a $\pm 20^\circ$ sector near the ecliptic plane. Histograms about the distribution of the ecliptic latitudes for objects from WISE and Pan-STARRS1 databases are presented in Fig. 3.4. We note that out of the two parameters, β is more sensitive to the observation geometries, and the violation of our assumption tends to result in the β solution being more inaccurate than the p solution.

3.2 Comparing populations

One of the key applications of our method is to detect differences between populations, and to analyze whether two samples belong to the same population or not.

Let S_1 and S_2 be two populations. We determine the shape elongation and spin latitude distributions for both populations. First, let us consider one of the populations. Let (p_i, β_j) be the grid points, $i = 1, \ldots, l$ and $j = 1, \ldots, m$, and let $\tilde{w}_{ij}$ be the occupation

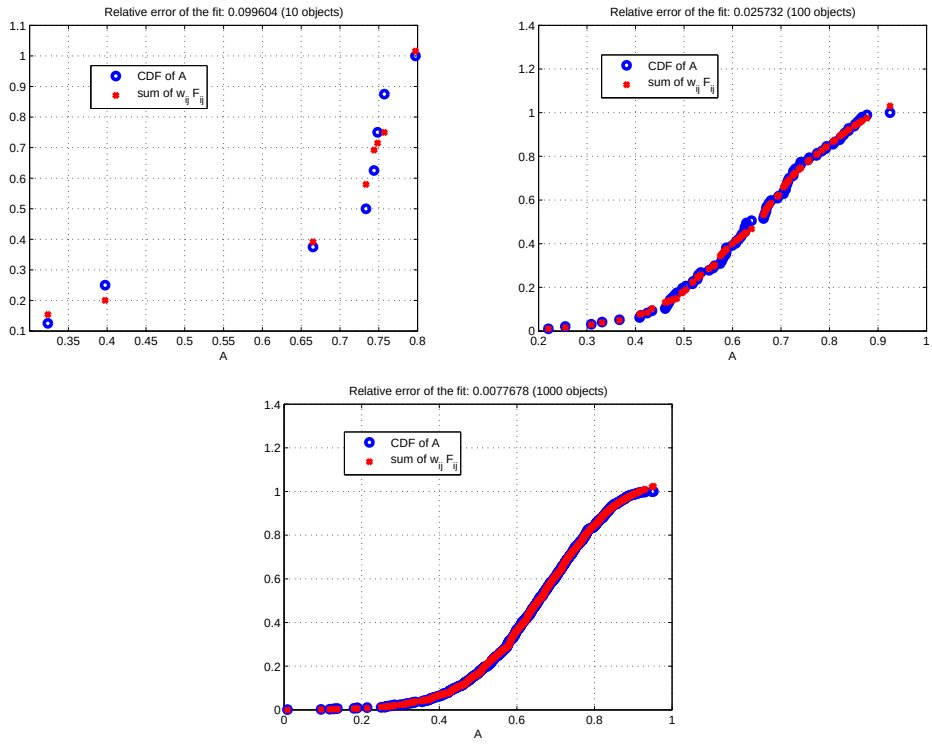

Figure 3.3: Plots of the function series $Mw = \sum_{ij} w_{ij} F_{ij}$ plotted versus the CDF $C = C(A)$ to illustrate how Mw converges to C as the number of asteroids grows, with sample populations of 10, 100 and 1000 asteroids.

numbers normalized such that $\sum_{i,j} \tilde{w}_{ij} = 1$. Now, the marginal density functions $f(p)$ and $f(\beta)$ can be computed for the population in the following way:

$$f(p_i) = \sum_{j=1}^{m} \tilde{w}_{ij}, \quad f(\beta_j) = \sum_{i=1}^{l} \tilde{w}_{ij}. \tag{3.16}$$

Next, let us compute the CDFs for the marginal DFs, denoting these CDFs as F_p and F_β:

$$F_{p_q} = \sum_{i=1}^{q} f(p_i), \quad F_{\beta_r} = \sum_{j=1}^{r} f(\beta_j). \tag{3.17}$$

With the CDFs $F_p(S_1)$, $F_p(S_2)$, $F_\beta(S_1)$ and $F_\beta(S_2)$ computed for both populations, we define the statistical differences between populations:

$$\begin{cases} D_p(S_1, S_2) = \alpha_k \left\| F_p(S_1) - F_p(S_2) \right\|_k \\ D_\beta(S_1, S_2) = \alpha_k \left\| F_\beta(S_1) - F_\beta(S_2) \right\|_k \end{cases}. \tag{3.18}$$

We typically compute the cases $k = 1$, $k = 2$ and $k = \infty$. The scaling coefficients α_k are arbitrary; we chose them to be $\alpha_1 = 1/4$, $\alpha_2 = 1$ and $\alpha_\infty = 2$ to fix the values of each statistical difference to the same magnitude. The case $k = \infty$ is the Kolmogorov–Smirnov

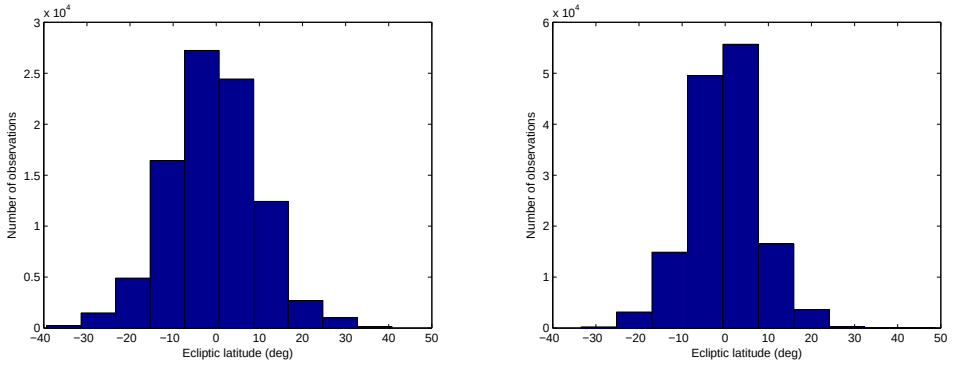

Figure 3.4: Histogram portraying the distribution of the ecliptic latitudes for targets from WISE (left) and Pan-STARRS1 (right) databases.

test typically used for statistical comparisons. However, one numerical value does not tell everything about the statistical differences between two populations, which is why we also compute the L^1 and L^2 differences. For example, the L^∞ norm detects if the statistical peaks are located in different positions, while the L^2 norm focuses on the average differences of populations. We note that the D_p and D_β values are not the same as the "p-value" typically used in statistics. As a general rule of thumb, the statistical difference between two populations tends to be significant if $D \gtrsim 0.2$. Nevertheless, we do not recommend blindly trusting three computed values, but also plotting the marginal DFs (and their CDFs) of the populations in the same figure, as a visual comparison is often more illustrative for analyzing the differences between populations.

3.3 Synthetic simulator

Unless we have literature references containing information about the shape elongation and spin latitudes for asteroid families available, the only way to evaluate the accuracy of the obtained (p, β) distribution is to run simulations with synthetic data. In the simulations, we use synthetic asteroid models where the p and β values of each asteroid is known, and we take geometries from the database we are testing, to determine how accurately the (p, β) distribution is reconstructed as the solution of the inverse problem from Sect. 3.1.3. The simulations can also be used for detecting systematic errors in the solution. The accuracy of the solution depends on the database used; we tested the accuracy of the WISE database in Nortunen and Kaasalainen (2017), and the accuracy of the Pan-STARRS1 database in Cibulková et al. (2018).

In order to avoid inverse crime (see Kaipio and Somersalo (2005)), it is important to create realistic simulations that are made with a model different from and independent of the analytical model used in the inversion. A further reason for the simulations is that ordinary error analysis cannot be used in our case since our actual observable is a cumulative distribution function rather than the original time series. CDFs do not have any well-defined pointwise errors as in time series: in this sense, fitting the model to the observations is different from the usual cases of goodness-of-fit and error levels.

We used DAMIT as the source for synthetic asteroid models, since it has a rich variety of realistic shapes. We applied basic transformations, such as stretching, on the DAMIT

models to obtain the desired shapes. For the ellipsoid model described in Sect. 3.1.3, we set a as the longest diameter in the equatorial xy-plane, and b as the width in the corresponding orthogonal direction. Then, the shape elongation is simply $p = b/a$.

We note that the error in the solution is dominated by the model noise. We also tried other methods for determining the shape elongation of an asteroid shape. For example, we computed ellipsoids that best fitted the contours of an asteroid shape. However, we observed no noticeable change in the accuracy of the solutions. Therefore, we opted to use the method described in the previous paragraph due to its simplicity and fast computation times.

Let us assume that the synthetic population has a fixed, single peak at (p^*, β^*), and the population contains N asteroids. Then, the following algorithm is repeated for each asteroid $i = 1, \ldots, N$:

1. We choose an asteroid model from DAMIT such that its shape elongation $p = b/a$ is close to the peak value p^* within a certain tolerance:

$$|p - p^*| \leq \epsilon. \tag{3.19}$$

   For example, we used a value of $\epsilon = 0.075$ to fix each elongation value to a distance of at most 0.15 p-units away from each other. In addition, we compute the normal and area for each facet of the asteroid model.

2. We choose a β from a Gaussian distribution, with the mean β^* and standard deviation σ (we used a value of $\sigma = 0.05$), and with a restriction $\beta \in (0, \pi/2)$. In addition, we fix the longitude λ by choosing it from a uniform distribution, $\lambda \in [0, 2\pi]$.

3. We extract geometries (direction vectors of the Sun and the Earth) from a data file belonging to the asteroid database we are studying, as well as the Julian dates of the measurements. If the geometries are not given in the data file directly, they need to be computed separately. The direction vectors should be transformed to the asteroid-fixed frame by using a coordinate transformation, (Kaasalainen et al., 2001), with the help of β and λ, which were fixed in the previous step. Let us denote the direction vectors in the asteroid's own frame $\mathbf{e}_{\text{sun}}$ and $\mathbf{e}_{\text{earth}}$, and let α be the phase angle between the Sun and the Earth. We filter out the cases where the condition

$$\alpha = \arccos(\mathbf{e}_{\text{sun}} \cdot \mathbf{e}_{\text{earth}}) \leq \alpha_{\text{tol}}. \tag{3.20}$$

   does not hold; depending on the database, α_{tol} should be at most 20 – 30 degrees. Finally, we compute the total brightness L for the synthetic asteroid, and add a minor Gaussian perturbation to L to simulate noise. For computing the total brightness of the synthetic DAMIT models, we use a combination of **Lommel–Seeliger and Lambert** scattering laws as in Kaasalainen and Lamberg (2006), since the models in DAMIT are mostly constructed using these scattering laws, and therefore we can reproduce the most realistic brightness values by using the same scattering laws. Alternatively, when using shape models from other sources than DAMIT, other scattering laws, such as **Hapke**, may also be used.

4. We proceed with the forward model as in Sect. 3.1.1. That is, we require at least five L values for a valid η estimate, and we may split the brightness data into multiple data sets to keep the changes in geometry small within each set. If the phase angle

changes too much within a set, we may apply phase correction on the set. Finally, we compute η and A for each set as in Eq. (3.2) and (3.7). With the amplitudes computed, we sort them in an increasing order. If we have k amplitudes in total, the CDF is simply $C(A_i) = i/k$, where $i = 1, \ldots, k$.

5. We compute the solution of the inverse problem as described in Sect. 3.1.3. A visual presentation of the actual joint distribution $\hat{f}(p, \beta)$ and the computed DF $f(p, \beta)$ is often illustrative to analyze the accuracy of the solution.

3.4 Conclusions

In this chapter, we have described our method of using distribution analysis to obtain information about a population when the data are scarce, and no model can be reconstructed for individual targets. In the forward model, we introduced our observable, brightness variation η, and the CDF $C(A)$ that we constructed from our observables; the connection between η and A is given in (3.7). The CDF can be expressed as a linear combination of analytical basis functions given in Sect. 3.1.2, and solving the inverse problem described in Sect. 3.1.3 gives us the occupation levels of each (p, β) bin. Thus, we obtain the joint distribution for the shape elongation p, and spin latitude β.

A solution with a rough-scale accuracy requires at least a 1000 observed targets in the population. Indeed, due to the poor quality of the data, a rough-scale solution is the best that can be obtained. Even so, our method turned out to be robust, and the accuracy of the solution has a high tolerance towards errors, often being unaffected by a moderate amount of noise. The total number of observations, which is high if the population is large enough, is what essentially dictates the accuracy of the solution, not the errors – in fact, the amount of noise in both the scarce data and the ellipsoid model is tremendous. The errors can be large, but as long as the population is large enough, we can obtain as much information from the scarce data as it is possible. While it is not possible to obtain a high-resolution solution, the method is still accurate enough to detect statistical peaks, as well as yield the approximate shape of the joint distribution. In a way, the distribution method is analogous to the famous telescope developed by **Galileo Galilei**, which could be used for detecting celestial bodies, but it could not obtain high-resolution images of the bodies. Similarly, our method can be used for discovering population-wide statistical peaks, and their approximate location in the (p, β) grid, but the method is not intended for high-resolution solutions. For other applications, the method can be used for discovering differences between the distributions of different populations, as explained in Sect. 3.2. While we have defined a few measures to compute the statistical differences between marginal distributions, it is usually more illustrative to perform a visual comparison between the marginal distributions, by plotting them in the same figure.

It is essential to note the scale for the number of targets (denoted as N on this paragraph) that can be observed with different data types. For the targets that can be accurately seen with a space probe, $N \propto 10^1$. For the targets that can be accurately modelled with the help of telescopes, adaptive optics or interferometry, $N \propto 10^2$. With dense photometry, about $N \propto 10^3$ targets can be modelled with good accuracy, and with sparse photometry, about $N \propto 10^4$ targets can be modelleted with rough accuracy. A more accurate listing of different observation techniques, the physical properties that can be derived, and the number of models that can be obtained, is given in Ďurech et al. (2015). In the case of scarce data, the number of targets that can be modelled on a population-wide level is

$N \propto 10^5$, or maybe even $N \propto 10^6$. It should be noted that the population-wide modelling is not something that we are comparing with other methods; it is essentially the only way to obtain information from scarce data.

In Sect. 3.3, we have written an algorithm for a synthetic simulator that can be used to test the accuracy and validity of our method with a given database. Each database behaves differently in terms of accuracy, systematic errors and tuning parameters, so rather than using our distribution analysis method blindly, the user should always run synthetic simulations on the database to acquire information on the accuracy levels of the database. In Chapter 4, we will discuss how we can use synthetic simulations to fine-tune our method to provide more accurate results with a given database.

4 Technical fine-tuning of the LEADER software package

We have developed the **LEADER** (Latitudes and Elongations of Asteroid Distributions Estimated Rapidly) software package for MATLAB to implement the method described in Chapter 3.1. The operating principles of the software package were explained in detail in Nortunen and Kaasalainen (2017), and the MATLAB code is available online[1].

The LEADER software package is roughly divided into three main components. The main algorithm implements the method described in Sect. 3.1. It analyzes each asteroid data file of the given database, and attempts to compute at least one brightness variation estimate η for each observed target, as explained in Sect. 3.1.1. Then the algorithm computes the joint distribution for p and β from the inverse problem explained in Sect. 3.1.3. Finally, the obtained solution is presented as a contour plot. Another component is the synthetic simulator, which attempts to reconstruct the joint (p, β) distribution to a synthetic population where the distribution is known in advance; the algorithm is explained in Sect. 3.3. The third component runs a comparison between populations: the main algorithm is carried out separately for two populations, and the obtained distributions are compared using the principles defined in Sect. 3.2.

As we noticed from Chapter 3, our method contains several adjustable parameters, such as the choice of (p, β) grid, and regularization parameters. Furthermore, when plotting the joint distributions obtained from the inverse problem, we can use a deconvolution filter to correct possible systematic errors in the solution. In this chapter, we will fine-tune the aforementioned settings to make sure the solution is as accurate as possible for most of the databases. The accuracy of the solution can be tested with synthetic simulations, where the actual (p, β) distribution of the population is known. We perform most of these simulations for the WISE database, as we have noticed that the parameter values that are applicable with WISE data typically work with other databases as well. However, we additionally perform some simulations with the Pan-STARRS1 database for completeness.

In addition to testing the WISE and Pan-STARRS1 databases with synthetic simulations, we performed simulations on synthetic data provided by **Josef Ďurech**. We applied phase correction on the synthetic observations, and used Hapke's scattering law to compute the brightness values for each target. Unfortunately, the noise level on the data was too large. As a result, the peak of the p distribution was placed far too much to the left, making the population seem more elongated than it actually was, and the β information was practically lost. The results of these simulations suggest that our method cannot recover information from overly noisy databases. Fortunately, there are asteroid databases where the noise level is small enough, and distributions can be reconstructed with good accuracy.

[1] https://github.com/northa7/LEADER (cited on April 30th, 2018)

4.1 Deconvolution

To acquire a good understanding of how the systematic error works when using WISE data, we performed a series of synthetic simulations with WISE-based synthetic data. We performed 50 runs of the simulator described in Sect. 3.3, each time with a different, randomized (p^*, β^*) peak. On each run, we used a population of 1000 synthetic asteroids. We reconstructed each (p, β) solution using the inverse problem, and located the peak of the solution distribution. We plotted the actual p peaks versus the computed p peaks, and similarly for β, in Fig. 4.1. The black dashed line is of the type "$y = x$", and it presents the case when the actual peak and the obtained peak coincide perfectly. For the shape elongation p, we perceived an obvious trend of the computed solution being about 0.1 units lower than the actual solution. A linear fit of the form "$y = x - 0.1$" is rather accurate, with the goodness of fit $R^2 \in [0.8, 0.9]$. The biggest error occurs when the actual p-peak is lower or equal to 0.4, which should be extremely rare for real asteroids. For β, however, the systematic error is harder to model. If the actual β peak value is low, the obtained peak shifts up, towards the middle values, while if the actual β peak is high, the obtained peak usually shifts down, towards middle values, but the solution distribution may alternatively spread towards the end value $\pi/2$, with the β peak appearing at the end of the distribution. A fit of the form "$y = \sqrt{x}$" or other similar power law would be possible for β, but we detected that such fit was inaccurate, with $R^2 \in [0.6, 0.7]$.

We attempted to correct the solution in the β direction by shifting the distribution, but the errors in β were so irregular that we omitted the shifting in order to avoid "false positives" in the correction.

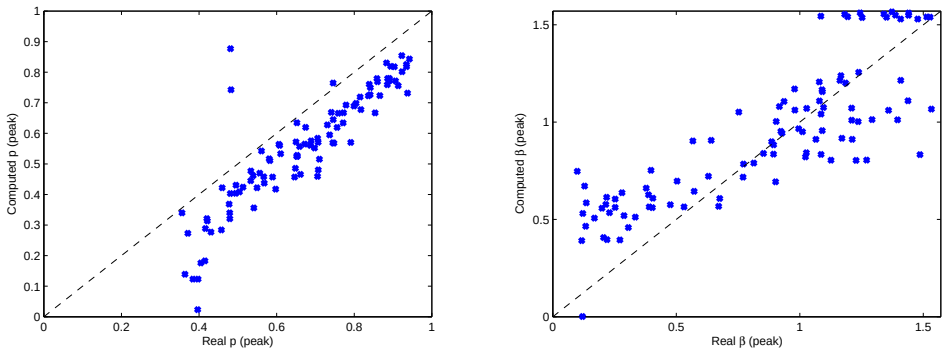

Figure 4.1: The actual p (left) and β (right) peak values plotted versus the computed p and β peak values with WISE-based synthetic models, with the limit curve of the form $y = x$ plotted in black to show the case when the computed values would be completely accurate.

In addition to testing how well the distribution peak is obtained, the synthetic simulator reveals how accurately the overall shape of the distribution is reconstructed. We observed that the reconstructed joint distribution tends to spread in both p and β directions when moving away from the peak. The spreading was systematic enough that it could be corrected by introducing dampening when moving away from the peak (assuming a one-peak solution). An example of the deconvolution filter applied to the computed solution for a synthetic population has been presented in Fig. 4.2. After applying the filter, the original "image" is reconstructed surprisingly accurately, excluding the shift in β direction.

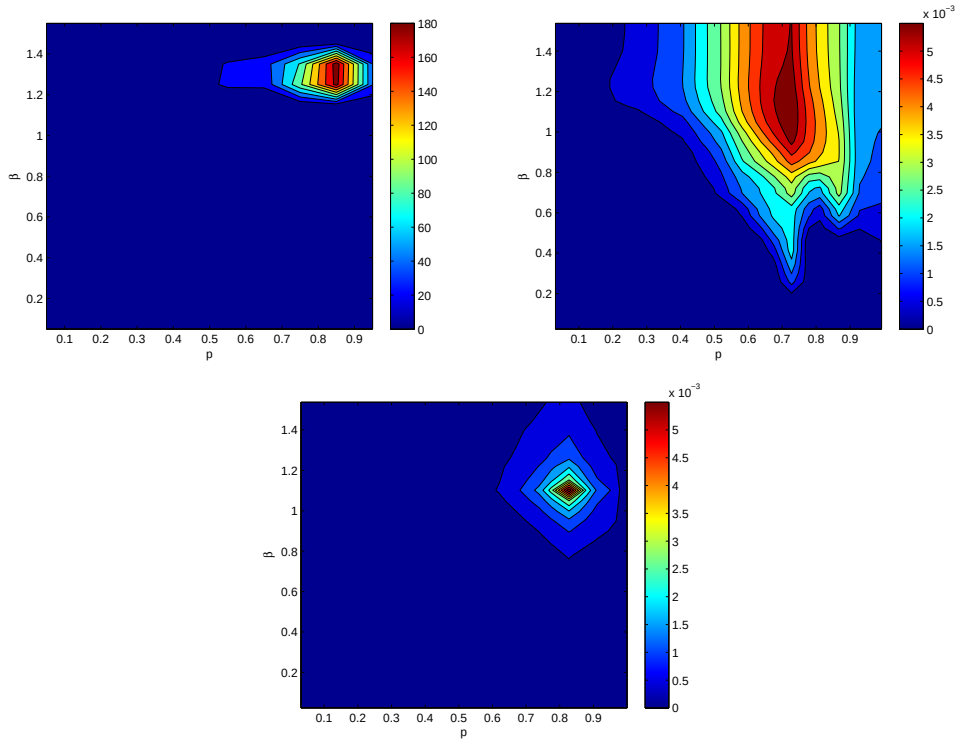

Figure 4.2: Actual joint distribution of (p, β) (top left) for a WISE-based synthetic population compared to the obtained solution (top right). On the bottom is the solution with deconvolution.

4.2 Choice of grid

We want to examine whether the choice of grid affects the accuracy of the solution. As some p and β values are less likely than some others, we may lower the resolution of those (p, β) bins, while keeping a large resolution with the bins that are more likely to be occupied. For p, the highly elongated values of $p < 0.5$ should be rare for realistic populations. Meanwhile for β, occupation numbers proportional to $\sin(\beta)$ would mean a uniform distribution on a sphere, so we expect small β values (perpendicular to the ecliptic plane) to be less common than β values close to $\pi/2$ (in the ecliptic plane). Therefore, on our experimental grid, we set p to have one bin in the intervals 0–0.25 and 0.25–0.4, and for values $p > 0.4$, we have one bin in the intervals 0.4–0.45, ..., 0.95–1. For β, we use one bin in the intervals 0°–18.5°, 18.5°–24.7°, 24.7°–31°, 31°–37.2° and 37.2°–43.4°, and for bigger angles, we use one bin in the intervals 43.4°–46.5°, ..., 86.9°–90°.

We used the same simulation setup as in Sect. 4.1 to test the accuracy of the solution with an equally spaced grid and the aforementioned selectively spaced grid. Plots of simulations with different grid choices for WISE and Pan-STARRS1 are presented in Figs. 4.3 and 4.4, respectively.

For WISE, it is more accurate to use equally spaced grids. If the resolution is small for low shape elongation values, the solution begins to get inaccurate when $p \lesssim 0.65$, shifting too much towards elongated values. For β, an equally spaced grid is more accurate, although it has a systematic error of shifting values from both extreme ends ($\beta = 0$ and $\beta = \pi/2$)

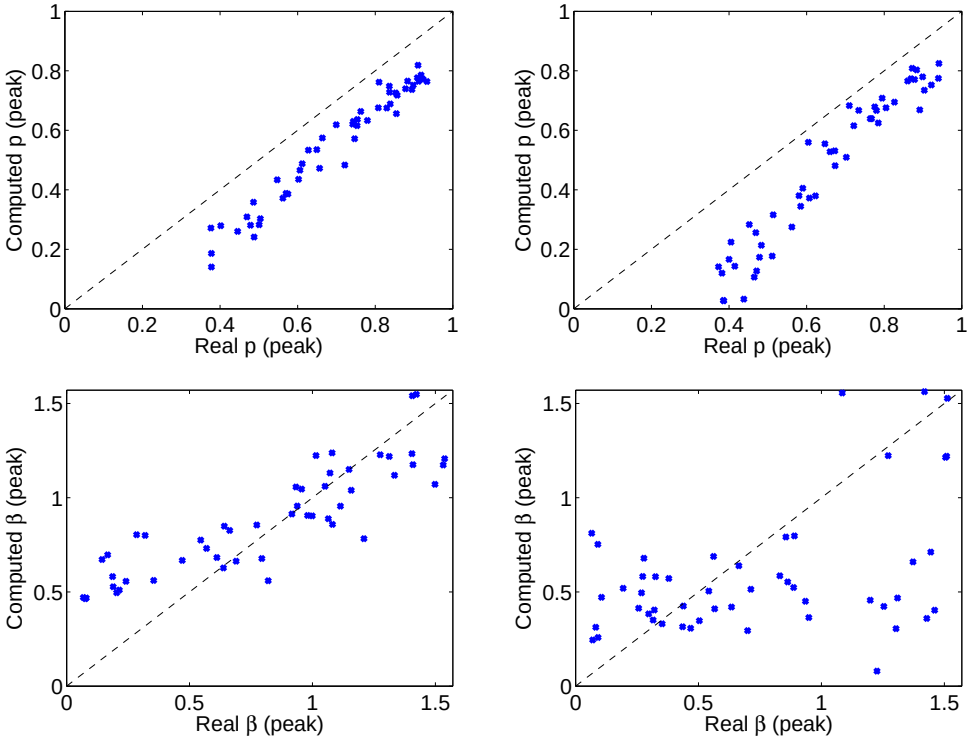

Figure 4.3: On top, the actual p peak values plotted versus the computed p peak values with WISE-based synthetic models, with the limit curve of the form $y = x$ plotted in black to show the case when the computed values would be completely accurate. Top left is for equally spaced bins, top right uses sparse spacing for smaller bins. On bottom, the same plots for β.

towards the middle. With a grid that uses a smaller resolution for values $\beta < \pi/4$, the solution clearly favors values that are perpendicular to the ecliptic plane, although it may invert the bodies in the ecliptic plane correctly.

For Pan-STARRS1, a grid with selective spacing yields more accurate results for β, while for equally spaced grids, the solution places all the peaks close to the ecliptic plane. With a selectively spaced grid, the p solution becomes less accurate when the peak is located at a value less than 0.5. For real applications, this is practically not an issue, since it is rare to encounter asteroid shapes with $p \ll 0.5$, let alone statistical peaks where $p < 0.5$. Therefore, we will use a selectively spaced (p, β) grid for Pan-STARRS1, as that way we can obtain at least some information about spin latitudes that are perpendicular to the ecliptic plane.

4.3 Regularization

Since we are using regularization in the inverse problem to smoothen and stabilize the solution, we should choose the regularization parameters δ_p and δ_β carefully. Too much regularization, that is, too high values for the regularization parameters, will eliminate much of the obtained information and provide only a flat solution. Meanwhile, if the regularization is insufficient, that is, too low values for the regularization parameters, the

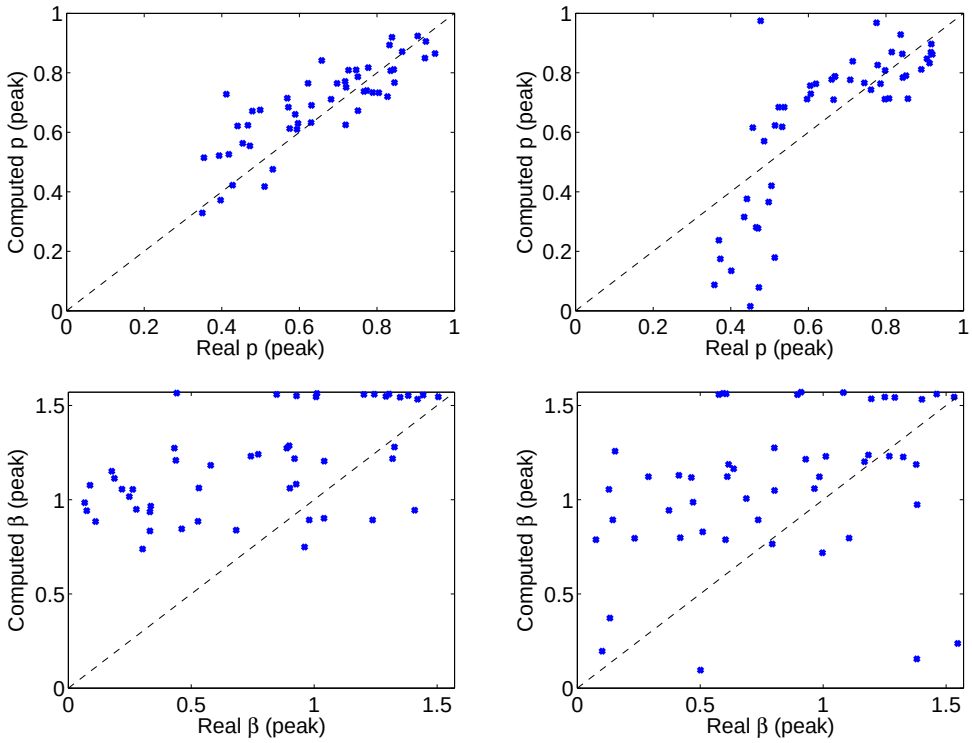

Figure 4.4: On top, the actual p peak values plotted versus the computed p peak values with Pan-STARRS1-based synthetic models, with the limit curve of the form $y = x$ plotted in black to show the case when the computed values would be completely accurate. Top left is for equally spaced bins, top right uses sparse spacing for smaller bins. On bottom, the same plots for β.

solution distribution will become unrealistically sharp, and it will additionally suffer from instability problems, being too sensitive to perturbations.

We use synthetic simulations to evaluate the accuracy of the solution when using different values for the regularization parameters δ_p and δ_β. The setup we used in the simulations is the same we used in Sect. 4.1. The results were plotted in Figs. 4.5 and 4.6 for δ_p and δ_β, respectively. For both regularization parameters, we tested how their values affect both p and β distributions, since p and β are not independent, and a change in the solution will thus affect both distributions.

We observed that sensible values for regularization parameters are generally $\delta_p \propto 10^{-1}$ and $\delta_\beta \propto 10^{0}$. Since β is more unstable than p, it is sensible that the β distribution requires more regularization than the p distribution. Hence, we have typically chosen simply $\delta_p = 0.1$ and $\delta_\beta = 1$, but these values can naturally be adjusted and fine-tuned according to the asteroid database used. For example, for WISE, we may choose slightly higher values for regularization parameters, such as $\delta_p \approx 0.25$ and $\delta_\beta \approx 3$, but for Pan-STARRS1, we noticed that either increasing or decreasing the values of the parameters from these values quickly causes the solution to deteriorate.

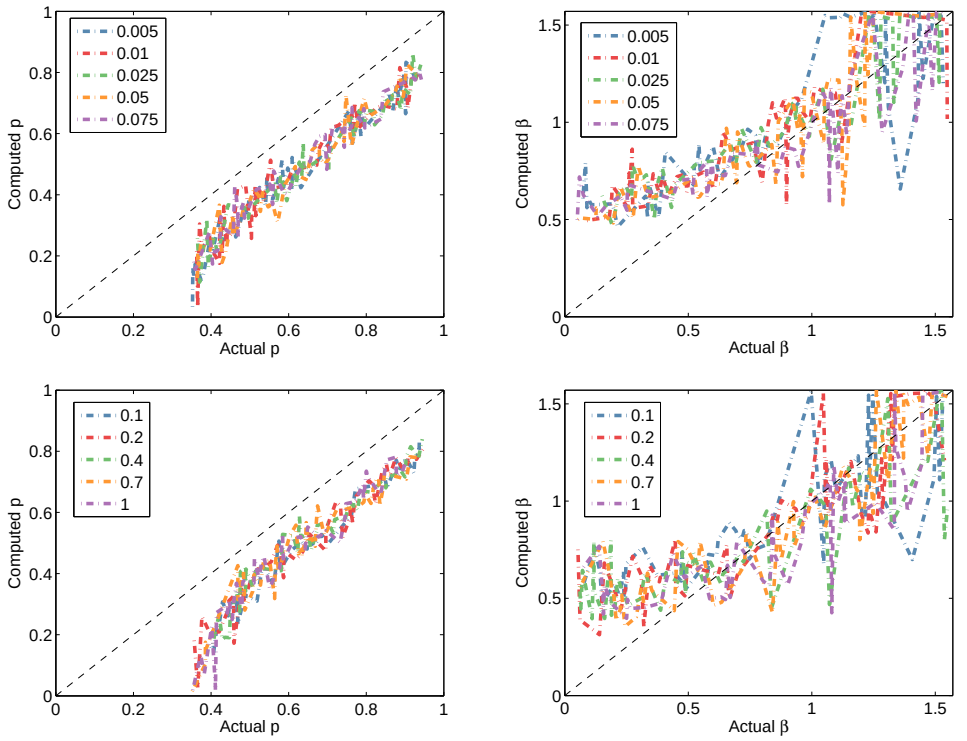

Figure 4.5: On left, the actual p peak values plotted versus the computed p peak values with WISE-based synthetic models with different values of δ_p. On bottom, the same plots for β with different values of δ_p.

4.4 Conclusions

We have used simulations with synthetic data in order to optimize the post-solution correction, grid setups and regularization parameters for the LEADER software package. We have noticed that some choices, such as the type of grid used, are highly dependent on the database. One type of a grid choice can improve the results considerably with one database, and significantly deteriorate the results with another database. The results presented in this chapter emphasize how the method of using distribution analysis to obtain information about a population should never be used "blindly", as a black box. Instead, the users of the software package should always evaluate the applicability of the method first, by performing simulations with the database that they aim to use.

The parameters we have discussed in this chapter have been tuned so that they can be applied to typical astronomical databases to give sensible results, but naturally, there are no global tuning parameters that give optimal results with every database. Instead, we encourage the users to try different sets of rules, both for data analysis and parameters, and see how the changes affect the accuracy of the solution. Adjusting these rules is simple, and the computation times for simulations are fast, so it is highly recommended that the user will fine-tune the settings of the software package so that it can be better tailored to work with the desired database.

The software package can also be used to experiment with different definitions of a

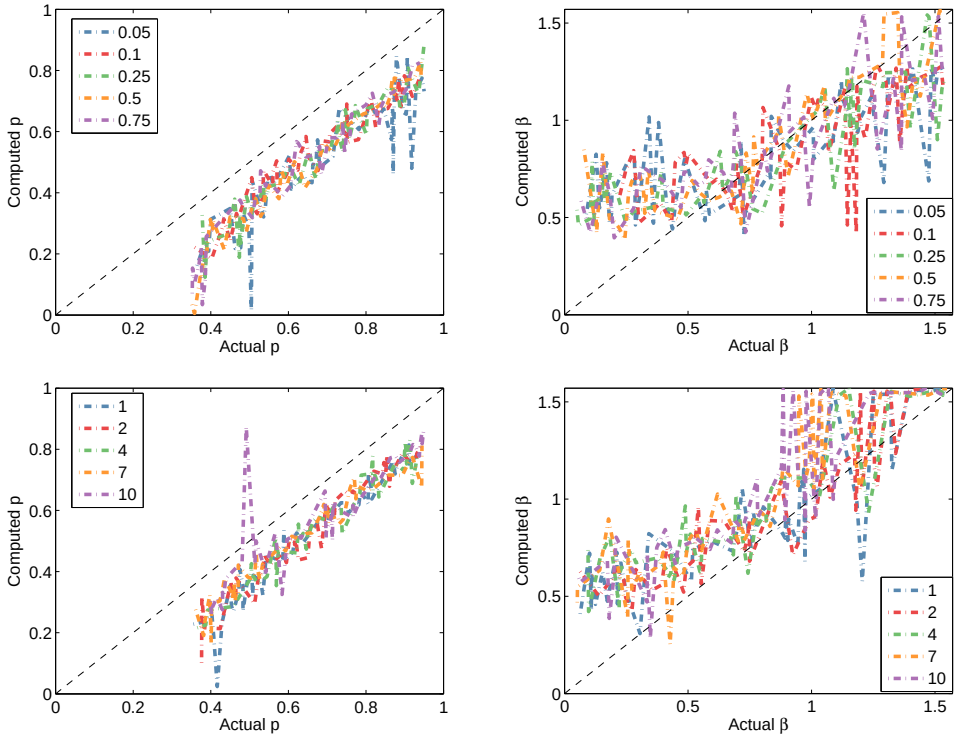

Figure 4.6: On left, the actual p peak values plotted versus the computed p peak values with WISE-based synthetic models with different values of δ_β. On bottom, the same plots for β with different values of δ_β.

population. While there are some astronomical ways of determining which targets belong to the same family, the definition of a population is ultimately arbitrary. If the user wants to study a group of asteroids with certain, shared properties, it is simple to create a script that filters the asteroids with the desired traits, and these asteroids form a new subpopulation. Another application for the software package is to test whether certain correlations exist, and similarly, to test whether a user-defined hypothesis holds or not. Some examples of testing correlations and hypotheses have been presented in Chapter 5. Some additional examples are presented in Nortunen et al. (2017) and Nortunen and Kaasalainen (2017). For instance, in Nortunen et al. (2017), we tested whether there is any correlation between the shape and size of the asteroids. We discovered that populations with a bigger diameter tend to have a higher concentration of spherical asteroids, and the spin axes are closer to the ecliptic plane.

5 Experiments with LEADER

5.1 Results from astronomical databases

In this section, we present some of the experiments we performed when comparing different asteroid populations. We focus on results that were not included in the publications Nortunen et al. (2017), Nortunen and Kaasalainen (2017) and Cibulková et al. (2018). On some of the experiments, we discovered that certain factors do not have a significant effect on the results. Nevertheless, we included these experiments in this dissertation, to notify what we have already studied.

As large asteroids are more spherical than small ones, we used LEADER to test if including large asteroids introduced any bias in our results. As an example, we plotted a comparison between the full Eos population and a sub-population with only small ($D < 20$ km) Eos asteroids (Figs. 5.1 and 5.2). We used data from the WISE database for both populations. For the p distribution, the inclusion of large asteroids mainly affects the width of the peak, and as the peak is always located in the middle of the distribution, there are no noticeable differences in the ratio of elongated and spherical bodies. The statistical differences are even smaller for the β distribution. As there seems to be no noticeable bias, we will take full populations and use them as such.

We performed a comparison between WISE and Pan-STARRS1 databases to examine what kind of differences there possibly are between the databases. In order to avoid bias related to the selection of targets, we considered only the asteroids that are found in both databases. There are about 70,000 mutual asteroids in both databases, which is more than sufficient for obtaining a smooth CDF $C(A)$. The comparisons of the solved p and β distributions have been plotted in Figs. 5.3 and 5.4, respectively. The p peak is located more towards spherical values for Pan-STARRS1, and the distribution is single-peaked. For WISE, the peak is unrealistically elongated; this is because the computed solution shifts the p values too much towards elongated values as seen in Fig. 4.1. Furthermore, the peak is very wide, and the distribution can even be considered double-peaked. As for the β distribution, the computed β peak lies closer to the ecliptic plane for Pan-STARRS1. From Fig. 4.4, we know that the computed solution tends to shift β too much towards the ecliptic plane for Pan-STARRS1.

We tested the effect of the orbit inclination I (the deviation of the orbit plane from that of the Earth) on the accuracy of the solution when using the Pan-STARRS1 database. In the simulations with synthetic data, we used subpopulations of 2000 asteroids for cases $\sin(I) \leq 0.2$ and $\sin(I) > 0.2$, with the geometries from Pan-STARRS1. The results have been plotted in Fig. 5.5. The differences in the accuracy of the computed p and β solutions are not significant for populations with low and high inclinations. Therefore, we conclude that the accuracy of our method is not affected by the inclination of orbits.

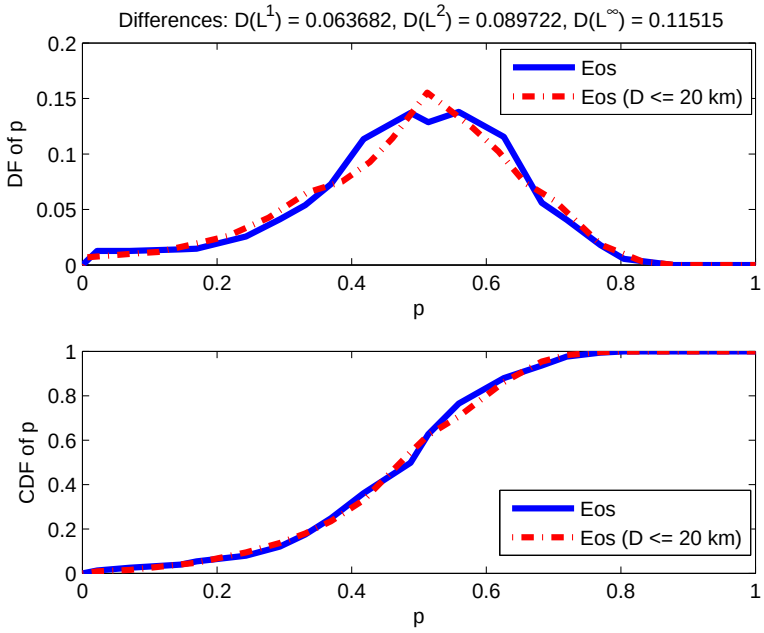

Figure 5.1: Comparison of the marginal DFs (top) of the shape elongation p for the whole Eos family and a subset of small (diameter $\leq$ 20 km) asteroids from the Eos family, and of their marginal CDFs (bottom). The asteroids were reconstructed by using measurements from WISE data.

5.2 The impact of YORP on the spin latitude

An intuitive guess would be that the spin latitude β is uniformly distributed. This was also the assumption used by Szabó and Kiss (2008), where the asteroid models had a randomly oriented spin axis, and the spin latitude was fixed. Indeed for large and spherical asteroid populations, the β distribution is almost uniform. As shown in Nortunen et al. (2017), large asteroids tend to be more spherical. In those cases, the YORP torques mostly vanish due to symmetry, and the effect of YORP is less prominent for large bodies. However, for small asteroids, which are typically more irregularly and asymmetrically shaped, the impact of YORP is significant, and causes β to be peaked, rather than isotropic. A β distribution proportional to $\sin(\beta)$ would mean that the spin latitude is uniformly distributed on a sphere, and the peak would be located at $\beta = \pi/2$. The effect of YORP shifts the β peak towards the pole of the ecliptic plane (Pravec and Harris, 2000; Rubincam, 2000).

Many astronomers have focused on studying the properties of certain asteroid families during the last decade. By performing a lightcurve analysis on observations from multiple epochs[1], it is possible to determine rotational poles, and carry out shape model fitting. Determining the pole for a large number of members in the asteroid family allows the astronomers to study the YORP evolution of the spin states of the family (Masiero et al., 2015). A significant result obtained from these lightcurve studies is that the YORP effect

[1]Here epoch means a point in time when certain geometries hold; the geomeries change as time passes.

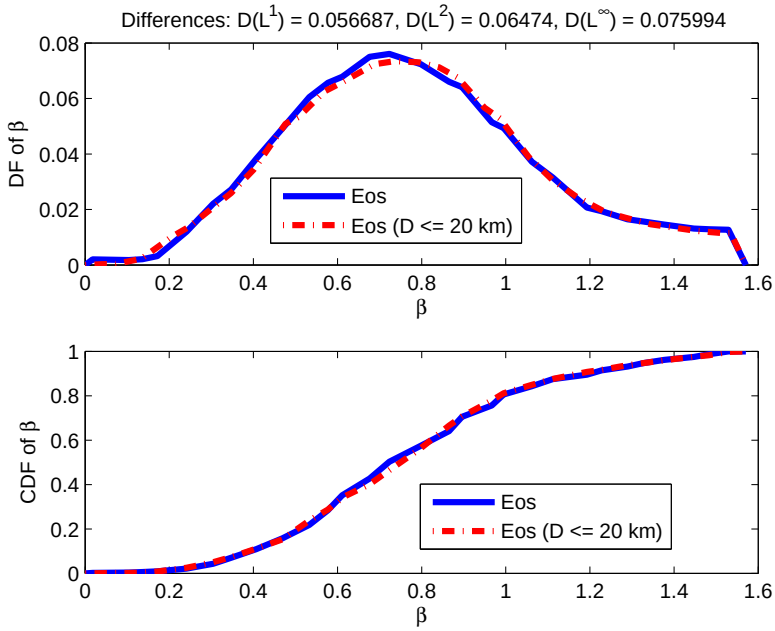

Figure 5.2: Comparison of the marginal DFs (top) of the spin latitude β for the Eos family and its subset (diameter $\leq$ 20 km), and of their marginal CDFs (bottom).

has been confirmed to affect the spin latitudes of bodies with similar ages (Vokrouhlický et al., 2003). The YORP effect not only creates a torque that alters the rotation rate of the asteroids, but it is also predicted that it rapidly alters the spin latitudes of small bodies towards being perpendicular to the ecliptic plane. That is, for populations with small bodies, some of the β values tend to cluster towards $\beta = 0$.

Several lightcurves of the **Koronis** family were measured by Slivan et al. (2003, 2008, 2009). Measurements from 30 members of the family showed a strongly anisotropic[2] distribution for the spin latitudes. The results were consistent with how the YORP effect was expected to alter rotation rates. For comparison, we considered all the asteroids in the Koronis family that were included in the WISE database, and computed their shape elongation and spin latitude distributions using the LEADER software package. The marginal distributions of p and β have been plotted in Fig. 5.6. No deconvolution has been used for the marginal DFs to preserve all of the information obtained in the solution. Both marginal DFs have been normalized so that $\sum_i f(p_i) = 1$ and $\sum_j f(\beta_j) = 1$. It is notable that such a large number of spin latitudes has been concentrated perpendicular to the ecliptic plane. Since the solution tends to avoid the extreme ends at $\beta = 0$ and $\beta = \pi/2$, it is possible that the actual β peak is even closer towards the pole. An additional sinusoidal curve has been plotted in the same figure to portray a uniform β distribution on a sphere, and spin latitudes with small β values (closer to the pole than the ecliptic plane) are clearly overrepresented. For comparison, Slivan et al. (2003) detected a large number of spin vectors clustered near the 20° (0.35 radians) and 45° (0.79 radians) angles for the Koronis family.

[2]Having different properties depending on the direction of measurement.

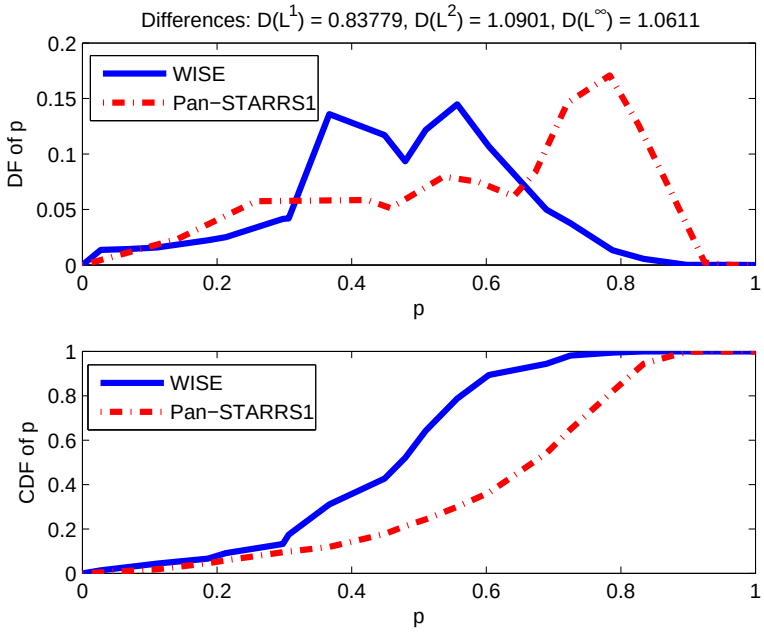

Figure 5.3: Comparison of the marginal DFs (top) of the shape elongation p for the WISE and Pan-STARRS1 databases, and of their marginal CDFs (bottom).

In this section, we have discussed the connection between the YORP effect and the spin latitudes of asteroids. The YORP effect tends to shift spin latitudes away from the ecliptic plane, towards β values closer to 0. This effect is especially noticeable for small, asymmetrical bodies. One of the first asteroid families where this connection was found was the Koronis family. We applied our method on the Koronis family and computed the p and β distributions of the population, and we confirmed that the β peak has indeed been shifted towards the ecliptic pole.

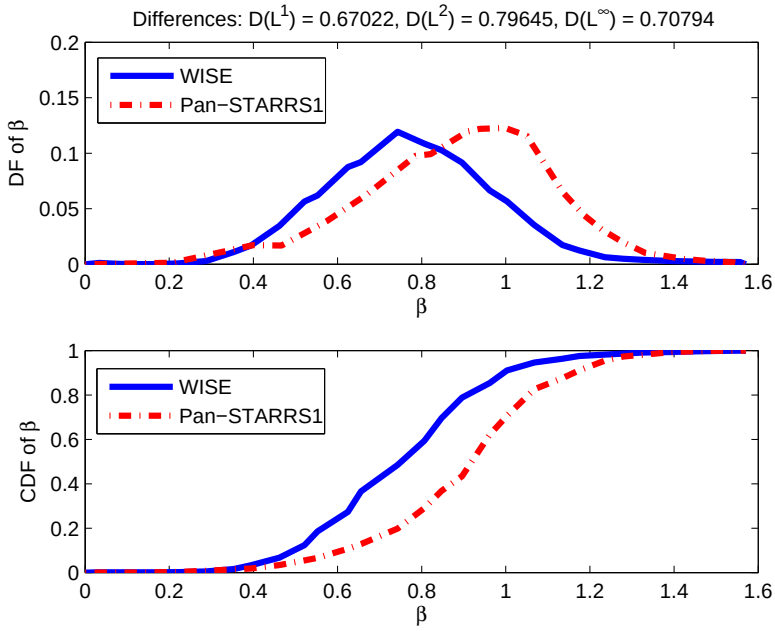

Figure 5.4: Comparison of the marginal DFs (top) of the spin latitude β for the WISE and Pan-STARRS1 databases, and of their marginal CDFs (bottom).

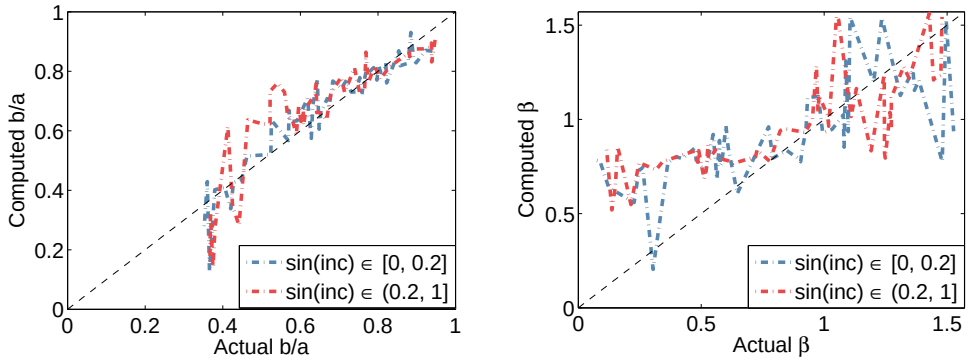

Figure 5.5: Simulations evaluating the accuracy of the p (left) and β (right) solutions with different orbit inclinations when using Pan-STARRS1 database.

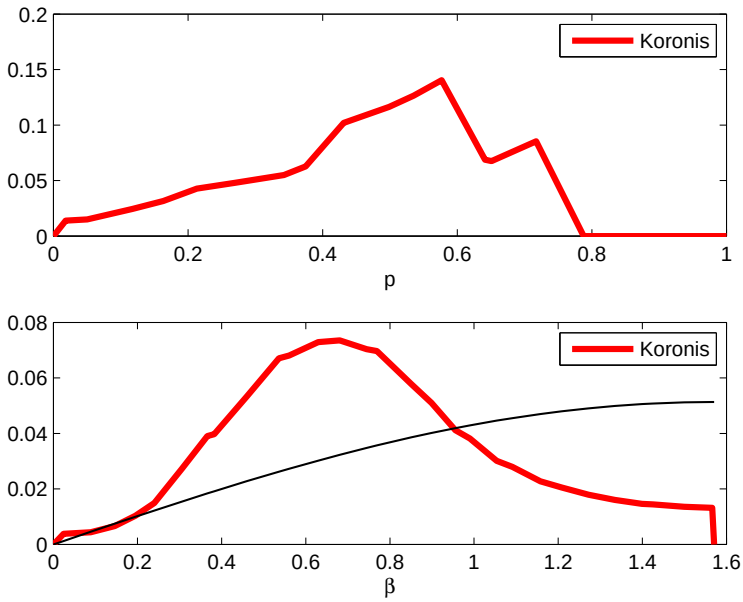

Figure 5.6: Marginal DFs of the shape elongation p (top) and the spin latitude β (bottom) for the Koronis family, using geometries from the WISE database. The black $\sin(\beta)$ curve depicts the situation when the β distribution is uniform on a sphere.

6 Conclusions and discussion

In this thesis, we have studied two new and innovative research problems. We have studied the YORP effect, where sunlight causes a torque that affects the rotation rate and obliquity of small celestial bodies, such as asteroids. Our focus was to derive an analytical theory on how the shape of the body affects the maximum effect of the light. While the YORP effect is fairly known and popular nowadays, the effect had not been examined analytically in detail before, so we aimed for a thorough mathematical study to derive the analytical expressions for the YORP torques. We also studied a complicated inverse problem, where we used brightness measurements to obtain information on the shape of asteroids. What made the situation problematic was that the available data were so weak that we could only obtain information on large asteroid populations, rather than individual targets. A few astronomers had attempted the same approach before, but without any mathematical theory or numerical validity tests to support them. In the field of inverse problems, the method is completely new.

When we derived the analytical torques for the YORP effect, we defined a new quantity: the YORP capacity $\mathcal{Y}$. The capacity portrays the maximum YORP effect on a body, scaled to be dimensionless and size-invariant. It is meant to describe the impact of shape on the magnitude of the YORP effect. We wanted to focus on the influence of shape alone, with no bias caused by the size of the object. We also estimated the upper bound of the YORP capacity, showing how the YORP capacity is theoretically unbounded for certain shapes, but practically, a finite upper bound exists.

We performed a stability analysis for the YORP effect to estimate the impact of errors in the asteroid models. With small perturbations in shape, the absolute change in $\mathcal{Y}$ remains bounded, so even for an unstable shape, we can usually obtain at least the correct order of magnitude for the YORP torques. The relative change in $\mathcal{Y}$, however, can be very large if the $\mathcal{Y}$ value is small. Depending on the magnitude of the YORP capacity, we divided shapes into different stability classes, where the shapes can be stable, semistable or unstable to shape perturbations. We studied the distribution of YORP capacities for the asteroid models found in DAMIT, and we noticed that a majority of the asteroids is YORP-unstable to global shape perturbations, while a fraction of the asteroids is YORP-semistable. Some asteroids, such as Itokawa, are so unstable that the observed YORP behavior changes completely when using different-resolution shape models. In such cases, only the order of magnitude can be obtained for the YORP torques, with the information on the actual torque behavior being unattainable. For future YORP research, this gives astronomers insight that obtaining accurate YORP information is not viable for certain shapes.

In the other case study, we examined an inversion problem where the data, given as time series, are so "scarce" that no model can be constructed for an individual target.

Rather than inspecting single targets, we focused on analyzing the statistical behavior of an object population. As far as the author knows, this approach has never been tried before to solve an inverse problem. Information on invididual targets is lost, but we discovered that we can still obtain a population-wide model. As our observable, we used a CDF that was constructed using brightness varation estimates from the brightness measurements. In Nortunen et al. (2017), we derived analytical basis functions, with their superposition yielding the observed CDF. With the help of these basis functions, we were able to unambiguously obtain the joint distribution for our two parameters, shape elongation p and spin latitude β. If the number of objects in the population is large, the observed CDF converges towards the actual CDF, and the joint distribution can be obtained with moderate accuracy. Even though the quality of the data and the ellipsoid model used to portray the observed asteroids are both weak, the distribution method turned out to be highly robust, as long as the population is large enough. A population of about 1000 targets or more is required for the observed CDF to converge.

The joint distribution for p and β is obtained on a rough-scale accuracy. Our distribution method is similar to the telescope developed by Galileo Galilei: it is applicable for making observations, but it cannot be used to acquire high-resolution "images". A population-level model is the best that can be reconstructed from scarce data, as the data are insufficient for providing information about individual targets. The modeling methods for dense or sparse photometries can be used for reconstructing $10^3 - 10^4$ targets, but for the rest, population-level modelling is the only way to obtain information about the targets.

The YORP effect has a prominent impact on the spin latitudes when considering populations that have plenty of small bodies. It causes the peak of the β distribution to shift away from the ecliptic plane, towards the ecliptic pole. The effect had been observed for the Koronis family before, and we confirmed the effect by using our distribution method to reconstruct the β distribution of the Koronis family. In Nortunen et al. (2017), we showed how the β peak is closer to the ecliptic plane for larger bodies. As a theoretical limit, the β distribution would approach the behavior of $\sin(\beta)$ for populations with only large and spherical bodies, as the YORP torques would either vanish or be insignificant for those bodies. The behavior of $\sin(\beta)$ would mean the spin latitudes are uniformly distributed on a sphere.

The properties of shape and spin distributions for different populations can be tested and compared with different populations using our distribution method. This leads to an interesting question about what defines a population, as the concept of a population is ultimately arbitrary. The LEADER software package allows the user to detect differences or possible correlations for different asteroid populations, as well as test different hypotheses about the properties of a population. In a way, the software package can be used as a data mining tool for examining whether there is a correlation between populations with certain characteristics.

For future research, astronomers may use the software package to experiment with different asteroid populations, as well as obtaining information on the properties of asteroid families. Some new asteroid databases will be available in the near future, such as the **Gaia**[1] survey in 2018, and **LSST**[2] (Large Synoptic Survey Telescope) in the late 2020s. The former contains data on tens of thousands of asteroids. Depending on the quality of the data, our distribution method may be the only way to obtain information about the shape

[1] http://sci.esa.int/gaia/ (cited on April 30th, 2018)
[2] https://www.lsst.org/ (cited on April 30th, 2018)

and spin properties of the asteroids surveyed on these databases. In addition, it may be possible to implement the distribution method on other fields with large populations, such as forest research.

A Appendix

A.1 Computing the spherical harmonics expansion

In a three-dimensional case, we can use **spherical harmonics** (Press et al., 2007) to express the radius $r(\theta, \phi)$, using an expansion

$$r(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_l^m Y_l^m(\theta, \phi), \tag{A.1}$$

assuming that the spherical harmonics expansion (A.1) converges and the shape is starlike (unique radius in each direction). The factors Y_l^m are Laplace's spherical harmonics, and they are of the form

$$Y_l^m(\theta, \phi) = P_l^m(\cos\phi)\mathrm{e}^{\mathrm{i}m\theta}, \tag{A.2}$$

where the terms $P_l^m = P_l^m(\cos\phi)$ are **associated Legendre polynomials** (Abramowitz and Stegun, 1965). We want to solve the coefficients $c_l^m \in \mathbb{C}$.

To avoid the use of complex numbers in computation, we can apply Euler's formula, writing $\mathrm{e}^{\mathrm{i}m\theta} = \cos(m\theta) + \mathrm{i}\sin(m\theta)$. As the radius $r(\theta, \phi)$ is a real-valued function, it is sufficient to consider the indices $m \geq 0$. Hence, the series expansion can be written as

$$r(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=0}^{l} \left[f_l^m \cos(m\theta) + g_l^m \sin(m\theta) \right] P_l^m(\cos\phi), \tag{A.3}$$

where $f_l^m, g_l^m \in \mathbb{R}$. During the actual computation, the spherical harmonics expansion must be approximated with a partial sum:

$$r(\theta, \phi) \approx \sum_{l=0}^{L} \sum_{m=0}^{l} \left[f_l^m \cos(m\theta) + g_l^m \sin(m\theta) \right] P_l^m(\cos\phi). \tag{A.4}$$

As we examine equation (A.3), we notice that the terms $\cos(m\theta)$ and $\sin(m\theta)$ are known, but it is yet to be determined how the terms $P_l^m(\cos\phi)$, f_l^m and g_l^m can be computed. First, consider the associated Legendre polynomials $P_l^m(\cos\phi)$. It is worth noting that the normalization constant frequently appearing in the expression of Y_l^m, as well as the Condon–Shortley phase factor $(-1)^m$ associated with the expression of P_l^m, have been omitted, as they can be included in the coefficients f_l^m and g_l^m. The associated Legendre polynomials can be presented as in Kaasalainen et al. (1992):

$$P_l^m(\cos\phi) = \sum_{n=0}^{l-m} A_{lm}^n \cos^n(\phi) \sin^m(\phi), \tag{A.5}$$

where the coefficient A_{lm}^n is defined recursively according to the algorithm below.

1. **If** $l > m$, **then**

$$A^n_{lm} = \frac{1}{l-m}\big[(2l-1)A^{n-1}_{l-1,\,m} - (l+m-1)A^n_{l-2,\,m}\big]. \tag{A.6}$$

2. **Else if** $l = m \neq 0$, **then** $n = 0$ and

$$A^0_{ll} = (2l-1)A^0_{l-1,\,l-1}. \tag{A.7}$$

3. **Else if** $l = m = 0$, **then** $n = 0$ and

$$A^0_{00} = 1. \tag{A.8}$$

4. **Else**,

$$A^n_{lm} = 0. \tag{A.9}$$

End.

Some of the associated Legendre polynomials have been tabulated in Belousov (1962). The expression $P^m_l(x)$ is commonly used in some sources, where $x \in [-1, 1]$. The notation can be made compatible with our convention by using a reparametrization with respect to the angle ϕ by setting $x = \cos(\phi)$.

Let us now derive expressions for the coefficients f^m_l and g^m_l. It can be shown that the associated Legendre polynomials are orthogonal:

$$\int_0^\pi P^m_l(\cos\phi)P^{m'}_{l'}(\cos\phi)\sin(\phi)\mathrm{d}\phi = \frac{2(l+m)!}{(2l+1)(l-m)!}\delta_{ll'}\delta_{mm'}, \tag{A.10}$$

where δ_{ij} is the Kronecker delta. We multiply both sides of equation (A.3) by the term $P^{m'}_{l'}(\cos\phi)\cos(m'\theta)$, where $m' \neq 0$, and integrating both sides results in

$$\begin{aligned}
&\int_0^{2\pi}\int_0^\pi r(\theta,\phi)P^{m'}_{l'}(\cos\phi)\cos(m'\theta)\sin(\phi)\mathrm{d}\phi\mathrm{d}\theta \\
=&\sum_{l=0}^{\infty}\sum_{m=0}^{l}\Big[f^m_l \underbrace{\int_0^{2\pi}\int_0^\pi P^m_l(\cos\phi)\cos(m\theta)P^{m'}_{l'}(\cos\phi)\cos(m'\theta)\sin(\phi)\mathrm{d}\phi\mathrm{d}\theta}_{:=I_1} \\
&+ g^m_l \underbrace{\int_0^{2\pi}\int_0^\pi P^m_l(\cos\phi)\sin(m\theta)P^{m'}_{l'}(\cos\phi)\cos(m'\theta)\sin(\phi)\mathrm{d}\phi\mathrm{d}\theta}_{:=I_2}\Big],
\end{aligned} \tag{A.11}$$

where

$$I_2 = \Big(\underbrace{\int_0^{2\pi}\sin(m\theta)\cos(m'\theta)\mathrm{d}\theta}_{=0}\Big)\Big(\int_0^\pi P^m_l(\cos\phi)P^{m'}_{l'}(\cos\phi)\sin(\phi)\mathrm{d}\phi\Big) = 0 \tag{A.12}$$

and

$$\begin{aligned}
I_1 &= \Big(\int_0^{2\pi}\cos(m\theta)\cos(m'\theta)\mathrm{d}\theta\Big)\Big(\int_0^\pi P^m_l(\cos\phi)P^{m'}_{l'}(\cos\phi)\sin(\phi)\mathrm{d}\phi\Big) \\
&\overset{(A.10)}{=} \frac{2(l+m)!}{(2l+1)(l-m)!}\delta_{ll'}\delta_{mm'}\int_0^{2\pi}\cos(m\theta)\cos(m'\theta)\mathrm{d}\theta \\
&= \frac{2\pi(l+m)!}{(2l+1)(l-m)!}\delta_{ll'}\delta_{mm'}.
\end{aligned} \tag{A.13}$$

Thus,

$$\begin{aligned}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\left[f_l^m I_1 + g_l^m I_2\right] &= \sum_{l=0}^{\infty}\sum_{m=0}^{l}\frac{2\pi(l+m)!}{(2l+1)(l-m)!}\delta_{ll'}\delta_{mm'}f_l^m \\ =&\frac{2\pi(l+m)!}{(2l+1)(l-m)!}f_l^m.\end{aligned} \tag{A.14}$$

Let us denote

$$N_l^m = \frac{2\pi(l+m)!}{(2l+1)(l-m)!}, \tag{A.15}$$

so the right-hand side can be written as $N_l^m f_l^m$, and the coefficient f_l^m will be

$$f_l^m = \frac{1}{N_l^m}\int_0^{2\pi}\int_0^{\pi} r(\theta,\phi)P_l^m(\cos\phi)\cos(m\theta)\sin(\phi)\mathrm{d}\phi\mathrm{d}\theta \tag{A.16a}$$

when $m \neq 0$. In the case where $m = 0$, the right-hand side of equation (A.11) becomes $2N_l^m f_l^m$, likewise, resulting in the following expression for the coefficient f_l^0:

$$f_l^0 = \frac{1}{2N_l^0}\int_0^{2\pi}\int_0^{\pi} r(\theta,\phi)P_l^0(\cos\phi)\sin(\phi)\mathrm{d}\phi\mathrm{d}\theta. \tag{A.16b}$$

Obtaining the expressions for the coefficients g_l^m is almost identical to deriving the expressions for f_l^m. We multiply both sides of equation (A.3) with the term $P_{l'}^{m'}(\cos\phi)\sin(m'\theta)$, where $m' \neq 0$, and integrate both sides. Using the notations from (A.11), now $I_1 = 0$ and $I_2 = N_l^m\delta_{ll'}\delta_{mm'}$, where N_l^m is the same as in equation (A.15). Therefore,

$$\sum_{l=0}^{\infty}\sum_{m=0}^{l}\left[f_l^m I_1 + g_l^m I_2\right] = \sum_{l=0}^{\infty}\sum_{m=0}^{l}\left[0 + N_l^m\delta_{ll'}\delta_{mm'}g_l^m\right] = N_l^m g_l^m, \tag{A.17}$$

and we obtain the coefficient g_l^m:

$$g_l^m = \frac{1}{N_l^m}\int_0^{2\pi}\int_0^{\pi} r(\theta,\phi)P_l^m(\cos\phi)\sin(m\theta)\sin(\phi)\mathrm{d}\phi\mathrm{d}\theta \tag{A.18a}$$

when $m \neq 0$. If $m = 0$, it holds that $g_l^0\sin(0\theta) = 0$ for all g_l^0 in equation (A.3). Then we can define, for example, that

$$g_l^0 = 0. \tag{A.18b}$$

Instead of obtaining the analytical solution, the coefficients f_l^m and g_l^m can also be computed using the least squares method. According to equation (A.4), the radius r can be expressed in a data point (θ_i, ϕ_j) as follows:

$$\begin{aligned}r(\theta_i,\phi_j) &\approx \sum_{l=0}^{L}\sum_{m=0}^{l} f_l^m P_l^m(\cos\phi_j)\cos(m\theta_i) + \sum_{l=0}^{L}\sum_{m=0}^{l} g_l^m P_l^m(\cos\phi_j)\sin(m\theta_j) \\ &\overset{g_l^0=0}{=} \sum_{l=0}^{L}\sum_{m=0}^{l} f_l^m P_l^m(\cos\phi_j)\cos(m\theta_i) + \sum_{l=1}^{L}\sum_{m=1}^{l} g_l^m P_l^m(\cos\phi_j)\sin(m\theta_j),\end{aligned} \tag{A.19}$$

where $i = 1, \ldots, N$ and $j = 1, \ldots, M$. Let us use a more compact notation ${}^{j}P_l^m = P_l^m(\cos\phi_j)$. We obtain the coefficients f_l^m and g_l^m by solving a system of linear equations

$$A\mathbf{f} = \mathbf{r}, \qquad A = \begin{bmatrix} A_C & A_S \end{bmatrix}, \tag{A.20a}$$

where

$$A_C = \left[\begin{array}{ccccc}
{}^{1}P_0^0\cos(0\theta_1) & {}^{1}P_1^0\cos(0\theta_1) & {}^{1}P_1^1\cos(1\theta_1) & \ldots & {}^{1}P_l^l\cos(l\theta_1) \\
{}^{2}P_0^0\cos(0\theta_1) & {}^{2}P_1^0\cos(0\theta_1) & {}^{2}P_1^1\cos(1\theta_1) & \ldots & {}^{2}P_l^l\cos(l\theta_1) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
{}^{M}P_0^0\cos(0\theta_1) & {}^{M}P_1^0\cos(0\theta_1) & {}^{M}P_1^1\cos(1\theta_1) & \ldots & {}^{M}P_l^l\cos(l\theta_1) \\
\hline
{}^{1}P_0^0\cos(0\theta_2) & {}^{1}P_1^0\cos(0\theta_2) & {}^{1}P_1^1\cos(1\theta_2) & \ldots & {}^{1}P_l^l\cos(l\theta_2) \\
{}^{2}P_0^0\cos(0\theta_2) & {}^{2}P_1^0\cos(0\theta_2) & {}^{2}P_1^1\cos(1\theta_2) & \ldots & {}^{2}P_l^l\cos(l\theta_2) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
{}^{M}P_0^0\cos(0\theta_2) & {}^{M}P_1^0\cos(0\theta_2) & {}^{M}P_1^1\cos(1\theta_2) & \ldots & {}^{M}P_l^l\cos(l\theta_2) \\
\hline
\vdots & \vdots & \vdots & \ddots & \vdots \\
\hline
{}^{1}P_0^0\cos(0\theta_N) & {}^{1}P_1^0\cos(0\theta_N) & {}^{1}P_1^1\cos(1\theta_N) & \ldots & {}^{1}P_l^l\cos(l\theta_N) \\
{}^{2}P_0^0\cos(0\theta_N) & {}^{2}P_1^0\cos(0\theta_N) & {}^{2}P_1^1\cos(1\theta_N) & \ldots & {}^{2}P_l^l\cos(l\theta_N) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
{}^{M}P_0^0\cos(0\theta_N) & {}^{M}P_1^0\cos(0\theta_N) & {}^{M}P_1^1\cos(1\theta_N) & \ldots & {}^{M}P_l^l\cos(l\theta_N)
\end{array}\right] \tag{A.20b}$$

and

$$A_S = \left[\begin{array}{ccccc}
{}^{1}P_1^1\sin(0\theta_1) & {}^{1}P_2^1\sin(0\theta_1) & {}^{1}P_2^2\sin(1\theta_1) & \ldots & {}^{1}P_l^l\sin(l\theta_1) \\
{}^{2}P_1^1\sin(0\theta_1) & {}^{2}P_2^1\sin(0\theta_1) & {}^{2}P_2^2\sin(1\theta_1) & \ldots & {}^{2}P_l^l\sin(l\theta_1) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
{}^{M}P_1^1\sin(0\theta_1) & {}^{M}P_2^1\sin(0\theta_1) & {}^{M}P_2^2\sin(1\theta_1) & \ldots & {}^{M}P_l^l\sin(l\theta_1) \\
\hline
{}^{1}P_1^1\sin(0\theta_2) & {}^{1}P_2^1\sin(0\theta_2) & {}^{1}P_2^2\sin(1\theta_2) & \ldots & {}^{1}P_l^l\sin(l\theta_2) \\
{}^{2}P_1^1\sin(0\theta_2) & {}^{2}P_2^1\sin(0\theta_2) & {}^{2}P_2^2\sin(1\theta_2) & \ldots & {}^{2}P_l^l\sin(l\theta_2) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
{}^{M}P_1^1\sin(0\theta_2) & {}^{M}P_2^1\sin(0\theta_2) & {}^{M}P_2^2\sin(1\theta_2) & \ldots & {}^{M}P_l^l\sin(l\theta_2) \\
\hline
\vdots & \vdots & \vdots & \ddots & \vdots \\
\hline
{}^{1}P_1^1\sin(0\theta_N) & {}^{1}P_2^1\sin(0\theta_N) & {}^{1}P_2^2\sin(1\theta_N) & \ldots & {}^{1}P_l^l\sin(l\theta_N) \\
{}^{2}P_1^1\sin(0\theta_N) & {}^{2}P_2^1\sin(0\theta_N) & {}^{2}P_2^2\sin(1\theta_N) & \ldots & {}^{2}P_l^l\sin(l\theta_N) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
{}^{M}P_1^1\sin(0\theta_N) & {}^{M}P_2^1\sin(0\theta_N) & {}^{M}P_2^2\sin(1\theta_N) & \ldots & {}^{M}P_l^l\sin(l\theta_N)
\end{array}\right]. \tag{A.20c}$$

For clarification, it holds for matrices A_C and A_S that $i = 1$ and $j = 1$ on the first row, $i = 1$ and $j = 2$ on the second row, and so forth, and $i = N$ and $j = M$ on the last row. Thus, matrices A_C and A_S have $N \cdot M$ rows in total. The matrix A_C has $\sum_{k=1}^{L+1} k$ columns,

and the matrix A_S has $\sum_{k=1}^{L} k$ columns. In addition,

$$\mathbf{f} = \begin{bmatrix} f_0^0 \\ f_1^0 \\ f_1^1 \\ \vdots \\ f_l^l \\ g_1^1 \\ g_2^1 \\ g_2^2 \\ \vdots \\ g_l^l \end{bmatrix} \quad \text{and} \quad \mathbf{r} = \begin{bmatrix} r(\theta_1, \phi_1) \\ r(\theta_1, \phi_2) \\ \vdots \\ r(\theta_1, \phi_M) \\ r(\theta_2, \phi_1) \\ r(\theta_2, \phi_2) \\ \vdots \\ r(\theta_2, \phi_M) \\ \vdots \\ r(\theta_N, \phi_1) \\ r(\theta_N, \phi_2) \\ \vdots \\ r(\theta_N, \phi_M) \end{bmatrix}. \tag{A.20d}$$

Both methods for obtaining the coefficients f_l^m and g_l^m have their advantages. The analytical method is effective if there are plenty of data points, and they create a grid (θ_i, ϕ_j), where $i = 1, \ldots, N$ and $j = 1, \ldots, M$. The least squares method is beneficial when the number of data points is relatively low – making it easier to invert the matrix – and the points do not appear in a grid, but as a function of the pair (θ_i, ϕ_i). For example, the asteroid database DAMIT contains asteroid shape models which have the set of vertices formed by using triangular tessellation. The vertices do not appear in a grid, but as data points $r(\theta_i, \phi_i)$, where $i = 1, \ldots, N$. Therefore, the least squares method is required for solving the coefficients for DAMIT shape models.

Bibliography

Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, ISBN 978-0486612720, 1965.

Aster, R. C., Borchers, B., and Thurber, C. H., *Parameter Estimation and Inverse Problems*. Elsevier Academic Press, 2nd Edition, 2012.

Belousov, S. L., "Tables of normalized associated Legendre polynomials," *Mathematical tables series*, vol. 18, p. 379, 1962.

Bottke, W. F., Vokrouhlický, D., Rubincam, D. P., and Brož, M., "Dynamical evolution of asteroids and meteoroids using the Yarkovsky effect," *Asteroids III*, pp. 395–408, 2002.

Bottke, W. F., Vokrouhlický, D., Rubincam, D. P., and Nesvorný, D., "The Yarkovsky and YORP effects: Implications for asteroid dynamics," *Annu. Rev. Earth Planet. Sci.*, vol. 34, pp. 157–191, 2006.

Breiter, S., Bartczak, P., Czekaj, M., Oczujda1, B., and Vokrouhlický, D., "The YORP effect on 25 143 Itokawa," *Astronomy and Astrophysics*, vol. 507, no. 2, pp. 1073–1081, 2009.

Cibulková, H., Nortunen, H., Ďurech, J., Kaasalainen, M., Vereš, P., Jedicke, R., Wainscoat, R. J., Mommert, M., Trilling, D. E., Schunová-Lilly, E., Magnier, E. A., Waters, C., and Flewelling, H., "Distribution of shape elongations of main belt asteroids derived from Pan-STARRS1 photometry," *Astronomy and Astrophysics*, vol. 611, p. A86, 2018.

Ďurech, J., Vokrouhlický, D., Kaasalainen, M., Weissman, P., Lowry, S. C., Beshore, E., Higgins, D., Krugly, Y. N., Shevchenko, V. G., Gaftonyuk, N. M., Choi, Y.-J., Kowalski, R. A., Larson, S., Warner, B. D., Marshalkina, A. L., Ibrahimov, M. A., Molotov, I. E., Michałowski, T., and Kitazato, K., "New photometric observations of asteroids (1862) Apollo and (25143) Itokawa – an analysis of YORP effect," *Astronomy and Astrophysics*, vol. 488, no. 1, pp. 345–350, 2008.

Ďurech, J., Kaasalainen, M., Warner, B. D., Fauerbach, M., Marks, S. A., Fauvaud, S., Fauvaud, M., Vugnon, J.-M., Pilcher, F., Bernasconi, L., and Behrend, R., "Asteroid models from combined sparse and dense photometric data," *Astronomy and Astrophysics*, vol. 493, pp. 291–297, 2009.

Ďurech, J., Carry, B., Delbo, M., Kaasalainen, M., and Viikinkoski, M., "Asteroid models from multiple data sources," *Asteroids IV*, pp. 183–202, 2015.

Ďurech, J., Hanuš, J., Alí-Lagoa, V., Delbo, M., and Oszkiewicz, D., "WISE data and sparse photometry used for shape reconstruction of asteroids," *Asteroids: New Observations, New Models, Proceedings of the International Astronomical Union*, vol. 10, no. S318, pp. 170 – 176, 2016.

Farinella, P., Vokrouhlický, D., and Hartmann, W. K., "Meteorite delivery via Yarkovsky orbital drift," *Icarus*, vol. 229, pp. 321 – 327, 1998.

Kaasalainen, M., "Physical models of large number of asteroids from calibrated photometry sparse in time," *Astronomy and Astrophysics*, vol. 422, pp. L39 – L42, 2004.

Kaasalainen, M. and Lamberg, L., "Inverse problems of generalized projection operators," *Inverse Problems*, vol. 22, pp. 749 – 769, 2006.

Kaasalainen, M. and Nortunen, H., "Compact YORP formulation and stability analysis," *Astronomy and Astrophysics*, vol. 558, p. A104, 2013.

Kaasalainen, M., Lamberg, L., Lumme, K., and Bowell, E., "Interpretation of lightcurves of atmosphereless bodies. I. General theory and new inversion schemes," *Astronomy and Astrophysics*, vol. 259, pp. 318 – 332, 1992.

Kaasalainen, M., Torppa, J., and Muinonen, K., "Optimization methods for asteroid lightcurve inversion II. The complete inverse problem," *Icarus*, vol. 153, pp. 37 – 51, 2001.

Kaipio, J. and Somersalo, E., *Statistical and Computational Inverse Problems*. Springer-Verlag New York, 2005.

Lawson, C. L. and Hanson, R. J., *Solving Least Squares Problems*. Prentice-Hall, 1974.

Masiero, J. R., DeMeo, F. E., Kasuga, T., and Parker, A. H., "Asteroid family physical properties," *Asteroids IV*, pp. 323 – 340, 2015.

McNeill, A., Fitzsimmons, A., Jedicke, R., Wainscoat, R., Denneau, L., Veres, P., Magnier, E., Chambers, K. C., Kaiser, N., and Waters, C., "Brightness variation distributions among main belt asteroids from sparse light curve sampling with Pan-STARRS 1," *Monthly Notices of the Royal Astronomical Society*, vol. 459, no. 3, pp. 2964 – 2972, 2016.

Miinalainen, T., "The YORP sensitivity of asteroids," Bachelor's thesis (in Finnish), Tampere University of Technology, 2014.

Nesvorný, D. and Vokrouhlický, D., "Analytic theory of the YORP effect for near-spherical objects," *The Astronomical Journal*, vol. 134, no. 5, pp. 1750 – 1768, 2007.

Nesvorný, D. and Vokrouhlický, D., "Analytic theory for the Yarkovsky-O'Keefe-Radzievski-Paddack effect on obliquity," *The Astronomical Journal*, vol. 136, no. 1, pp. 291 – 299, 2008.

Nortunen, H. and Kaasalainen, M., "LEADER: fast estimates of asteroid shape elongation and spin latitude distributions from scarce photometry," *Astronomy and Astrophysics*, vol. 608, p. A91, 2017.

Nortunen, H., Kaasalainen, M., Ďurech, J., Cibulková, H., Alí-Lagoa, V., and Hanuš, J., "Shape and spin distributions of asteroid populations from brightness variation estimates and large databases," *Astronomy and Astrophysics*, vol. 601, p. A139, 2017.

Paddack, S. J., "Rotational bursting of small celestial bodies: Effects of radiation pressure," *J. Geophys. Res.*, vol. 74, pp. 4379 – 4381, 1969.

Paddack, S. J. and Rhee, J. W., "Rotational bursting of interplanetary dust particles," *Geophys. Res. Lett.*, vol. 2, pp. 365 – 367, 1975.

Pravec, P. and Harris, A. W., "Fast and slow rotation of asteroids," *Icarus*, vol. 148, pp. 12 – 20, 2000.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes: The Art of Scientific Computing.* 3rd edition, New York: Cambridge University Press, ISBN 978-0-521-88068-8, 2007.

Radzievskii, V. V., "The influence of anisotropy of reemitted sunlight on the orbital motion of asteroids and meteoroids," *Astron. Zh.*, vol. 29, pp. 162 – 170, 1952.

Radzievskii, V. V., "A mechanism for the disintegration of asteroids and meteorites," *Dokl. Akad. Nauk SSSR*, vol. 97, pp. 49 – 52, 1954.

Rubincam, D. P., "Asteroid orbit evolution due to thermal drag," *J. Geophys. Res.*, vol. 100, pp. 1585 – 1594, 1995.

Rubincam, D. P., "Yarkovsky thermal drag on small asteroids and Mars-Earth delivery," *J. Geophys. Res.*, vol. 103, pp. 1725 – 1732, 1998.

Rubincam, D. P., "Radiative spin-up and spin-down of small asteroids," *Icarus*, vol. 148, pp. 2 – 11, 2000.

Scheeres, D. J., Abe, M., Yoshikawa, M., Nakamura, R., Gaskell, R. W., and Abell, P. A., "The effect of YORP on Itokawa," *Icarus*, vol. 188, pp. 425 – 429, 2007.

Slivan, S. M., Binzel, R. P., Crespo da Silva, L. D., Kaasalainen, M., Lyndaker, M. M., and Krčo, M., "Spin vectors in the Koronis family: comprehensive results from two independent analyses of 213 rotation lightcurves," *Icarus*, vol. 162, pp. 285 – 307, 2003.

Slivan, S. M., Binzel, R. P., Boroumand, S. C., Pan, M. W., Simpson, C. M., Tanabe, J. T., Villastrigo, R. M., Yen, L. L., Ditteon, R. P., Pray, D. P., and Stephens, R. D., "Rotation rates in the Koronis family, complete to H≈11.2," *Icarus*, vol. 195, pp. 226 – 276, 2008.

Slivan, S. M., Binzel, R. P., Kaasalainen, M., Hock, A. N., Klesman, A. J., Eckelman, L. J., and Stephens, R. D., "Spin vectors in the Koronis family. II. Additional clustered spins, and one stray," *Icarus*, vol. 200, pp. 514 – 530, 2009.

Szabó, G. and Kiss, L., "The shape distribution of asteroid families: Evidence for evolution driven by small impacts," *Icarus*, vol. 196, no. 1, pp. 135 – 143, 2008.

Vokrouhlický, D., Nesvorný, D., and Bottke, W. F., "The vector alignments of asteroid spins by thermal torques," *Nature*, vol. 425, pp. 147 – 151, 2003.

Vokrouhlický, D., Bottke, W. F., Chesley, S. R., Scheeres, D. J., and Statler, T. S., "The Yarkovsky and YORP effects," *Asteroids IV*, pp. 509 – 531, 2015.

Yarkovsky, I. O., "The density of luminiferous ether and the resistance it offers to motion (in Russian)," *published privately by the author in Bryansk*, 1901.

Öpik, E. J., "Collision probabilities with the planets and the distribution of interplanetary matter," *Proc. R. Irish Acad.*, vol. A54, pp. 165 – 199, 1951.

Publications

Publication I

Kaasalainen, M. and Nortunen, H., "Compact YORP formulation and stability analysis," *Astronomy and Astrophysics*.

A&A 558, A104 (2013)
DOI: 10.1051/0004-6361/201322221

Astronomy & Astrophysics

Compact YORP formulation and stability analysis

Mikko Kaasalainen and Hari Nortunen

Department of Mathematics, Tampere University of Technology, PO Box 553, 33101 Tampere, Finland
e-mail: mikko.kaasalainen@tut.fi

Received 6 July 2013 / Accepted 3 September 2013

ABSTRACT

We present a concise analytical formulation of the YORP effect, with exact formulae for torques on convex bodies and motion-averaged components applicable to any shapes. We analyze the main features of the secular torques for zero and nonzero thermal inertia that are function series dependent on only a few coefficients. Using these, we investigate the stability of the YORP effect against shape perturbations with analytical and numerical estimates. We define a quantity describing the YORP capacity of any shape, and estimate YORP stability with it.

Key words. methods: analytical – methods: numerical – radiation mechanisms: thermal – minor planets, asteroids: general – minor planets, asteroids: individual: (1862) Apollo – minor planets, asteroids: individual: (3103) Eger

1. Introduction

The effect of sunlight on the dynamics of small asteroids is now known to be considerable over long time spans. The anisotropic reflection and thermal re-emission of solar photons is essentially equivalent to a large number of small rocket engines distributed on the surface of an asteroid. Thus the dynamical mechanisms are usually divided into two categories: the net propulsion force changing the orbital motion (the Yarkovsky effect), and the net torque affecting the rotational dynamics (the Yarkovsky-O'Keefe-Radzievskii-Paddack or YORP effect). Both have been observed directly (Chesley 2003; Kaasalainen et al. 2007; Lowry et al. 2007; Ďurech et al. 2008, 2012), and there is clear indirect evidence of their long-term role in the evolution of asteroid orbits and spin states (Bottke et al. 2001; Vokrouhlický et al. 2003, 2006).

The analytical modeling of the YORP torques, or any other illumination-dependent functions, on general shapes is constrained by the concept of insolation: in which geometries is a surface patch visible? For convex bodies insolation is trivial, and they are the only ones for which exact analytical torques are possible. Various semianalytical approaches have developed small nonconvex perturbations while neglecting numerical ray-tracing within some perturbation size (e.g., Breiter & Michalska 2008; Nesvorný & Vokrouhlický 2007, 2008a; hereafter NV07 and NV08). Unfortunately, the shadowing error cannot be estimated in terms of the perturbation parameter, so no validity regime can be established for such approximations even when the perturbation is very small.

We present the fundamentals of the YORP effect and its stability and symmetry properties in a conceptually and technically simple analytical and exact formulation. For example, it is easy to identify the enigmatic YORP instability with the "trash coefficients" of a Laplace series, and to show that, while quadrant-symmetric bodies have no secular YORP torques for zero thermal inertia ($K = 0$), they acquire a nonvanishing component for $K \neq 0$. Breiter et al. (2011; hereafter BRV11) arrive at some of the results of Sect. 3 ($K = 0$) by a different context and approach. We provide easy-to-use formulae directly applicable to typical body representations such as polyhedra. The motion-averaged formulation pertains to any shapes.

We discuss the basic concepts and functions in Sect. 2. The cases of zero and nonzero thermal inertia are studied in, respectively, Sects. 3 and 4. The YORP effect can be very sensitive to the details of the shape and other physical characteristics of the body (Statler 2009; Rozitis & Green 2012), so we discuss YORP stability in Sect. 5. We sum up in Sect. 6, and numerical formulae for convex polyhedral shape representations are presented in an appendix.

2. Problem setup and definitions

Let the direction of the illumination source in a coordinate system fixed to the target body (the z-axis coinciding with the rotation axis) be denoted by $\omega \in S^2$. Here entities on the unit sphere S^2, defined by two direction angles, are identified with unit vectors in $\mathbb{R}^3$. Thus, e.g., the outward unit normal vector $\eta \in S^2$ is given by $\eta = \eta(\vartheta, \psi)$ (with ϑ measured from the pole), $0 \le \vartheta \le \pi$, $0 \le \vartheta < 2\pi$; i.e.,

$$\eta_1 = \sin\vartheta\cos\psi, \quad \eta_2 = \sin\vartheta\sin\psi, \quad \eta_3 = \cos\vartheta. \tag{1}$$

Likewise, $\omega = \omega(\theta, \varphi)$ (the subsolar coordinates). The position of the asteroid around the Sun in the orbital frame is given by λ, $0 \le \lambda < 2\pi$, and the tilt (obliquity) of the rotation axis from the orbital pole is denoted by ϵ, $0 \le \epsilon \le \pi$. The orbital frame and the pole direction of the asteroid are chosen such that

$$\cos\theta = (\cos\lambda, \sin\lambda, 0)\cdot(\sin\epsilon, 0, \cos\epsilon) = \cos\lambda\sin\epsilon. \tag{2}$$

For clarity, we assume a circular orbit around the Sun. Ellipticity is straightforward to include as described in NV07 and NV08, and in any case the modifications due to ellipticity are only of order $(\sqrt{1-e^2})^{-1}$ for eccentricity e (NV07). Denoting the rotation angle of the asteroid around its axis by $\varphi' := \Omega t$, the transformation of a vector $\boldsymbol{v}$ from the orbital frame to the body frame reads

$$\boldsymbol{v}_{\rm ast} = \mathbf{R}_z(\varphi')\mathbf{R}_y(\epsilon)\boldsymbol{v}_{\rm orb}, \tag{3}$$

where $\mathbf{R}_i$ are the usal rotation matrices for z- and y-axes (Kaasalainen & Lamberg 2006). The transformations to an oblique frame aligned with the rotation axis but not rotating with the asteroid are

$$\boldsymbol{v}_\epsilon = \mathbf{R}_z(-\varphi')\boldsymbol{v}_{\rm ast} = \mathbf{R}_y(\epsilon)\boldsymbol{v}_{\rm orb}. \tag{4}$$

The oblique frame is needed in computing the torque component that changes the obliquity of the spin axis. We assume here that λ remains virtually constant during $2\pi/\Omega$.

2.1. Characteristic torque function

The thermal YORP torque by emitted photons is given by (NV07, NV08)

$$\tau = -\frac{2\gamma}{3v_c}\int_{\mathcal{B}} (\boldsymbol{x}\times\eta)T^4\,\mathrm{d}S, \tag{5}$$

where the integral is taken over the whole surface of the body $\mathcal{B}$ (of arbitrary shape and topology), T is the surface temperature, γ is the product of the emissivity of the material in thermal wavelengths and the Stefan-Boltzmann constant, and v_c is the speed of light. (As shown in Nesvorný & Vokrouhlický 2008b, the torque due to absorbed photons vanishes identically when averaged over orbit revolution and body rotation.)

We define the characteristic torque function $\mathcal{T}$ of a body as the three-dimensional vector

$$\mathcal{T}(\xi) := \boldsymbol{x}(\xi)\times\eta(\xi), \tag{6}$$

where ξ is any two-dimensional parametrization of the location on the surface. For a convex body,

$$\mathrm{d}S = G(\eta)\,\mathrm{d}\sigma, \quad \mathrm{d}\sigma = \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\psi, \tag{7}$$

where $G(\eta)$ is the curvature function, so we consider $\mathcal{T}$ on S^2:

$$\mathcal{T}(\eta) := G(\eta)\boldsymbol{x}(\eta)\times\eta \tag{8}$$

(cf. BRV11). For asteroids, $G(\eta)$ is typically an initial product from observational data, and $\boldsymbol{x}(\eta)$ is derived from it as the solution of the Minkowski problem (usually as a convex polyhedron; Kaasalainen et al. 1992, 2006).

Expanding $\mathcal{T}$ as a Laplace series, we have

$$\mathcal{T}(\eta) = \sum_{l\ge 0}\sum_{m=-l}^{l} \mathbf{T}_{lm}\,Y_l^m(\vartheta,\psi). \tag{9}$$

Here we ignore the customary normalization constants of the spherical harmonics Y_l^m as they can always be introduced as normalization factors when actually taking orthogonality integrals, so $Y_l^m(\vartheta,\psi) = P_l^m(\cos\vartheta)e^{im\psi}$. Written explicitly as a real-valued function,

$$\mathcal{T}_i(\eta) = \sum_{l\ge 0}\sum_{m=0}^{l} P_l^m(\vartheta,\psi)\left[T_{lm}^{i,\cos}\cos(m\psi) + T_{lm}^{i,\sin}\sin(m\psi)\right]. \tag{10}$$

The possible symmetries of $\mathcal{T}$ are evident in $\mathbf{T}_{lm}$. For example, geometric inspection shows that if the body is symmetric about a line in the xy-plane, $\mathbf{T}_{l0} \equiv 0$. If the body is quadrant-symmetric (e.g., an ellipsoid), the Fourier terms are $\mathcal{T}_1 \sim \sin m\psi$, $\mathcal{T}_2 \sim \cos m\psi$ with m odd, while $\mathcal{T}_3 \sim \sin m\psi$ with m even. The coefficients $\mathbf{T}_{lm}$ are readily calculated with the formulae in the appendix.

3. YORP torque and its components for $K = 0$

When there is no thermal inertia; i.e., the thermal conductivity $K = 0$, the absorbed flux is emitted immediately. Denoting

$$\mu := \eta\cdot\omega, \tag{11}$$

the immediate emission means that, for a surface patch $\mathrm{d}S$,

$$\gamma T^4 = (1-A)F_o\mu, \tag{12}$$

where A is the surface albedo and F_o the incoming solar flux. Thus we have, denoting the illuminated portion of the surface by A_+ and writing the convex version explicitly,

$$\tau = Q\int_{A_+}\mu\,\mathcal{T}(\xi)\,dS = Q\int_{\mu\ge 0}\mu\,\mathcal{T}(\eta)\,\mathrm{d}\sigma, \tag{13}$$

where Q has the physical dimension of pressure:

$$Q = -\frac{2}{3v_c}(1-A)F_o. \tag{14}$$

The effect of albedo variegation $A(\eta)$ on the surface can be absorbed in $\mathcal{T}(\eta)$ by replacing $G(\eta)$ with $[1-A(\eta)]G(\eta)$ in (8) and defining $Q = -2F_0/3v_c$.

The YORP torque in the body frame is now obtained in the same way as the opposition brightness for convex bodies (Kaasalainen et al. 1992; now we just have the coefficients of a three-dimensional Laplace series). Thus we have the fundamental result for the YORP torque τ:

$$\tau(\theta,\varphi) = Q\sum_{lm}\mathbf{t}_{lm}\,Y_l^m(\theta,\varphi), \tag{15}$$

where, for convex bodies, $\mathbf{t}_{lm} = 2\pi k_l\,\mathbf{T}_{lm}$, and

$$k_l = \frac{1}{l+2}(-1)^{l/2-1}\frac{(l-3)!!}{l!!}, \quad l\ge 2,\ l\,\text{even}. \tag{16}$$

Other values of l are not of interest as the odd ones drop out for all shapes in averaging (see below), and $\mathbf{T}_{00}$ vanishes identically due to Gauss' theorem.

As pointed out in BRV11, the torque τ for any shape, computed on a grid of (θ,φ) by ray-tracing, can, of course, be expressed as a Laplace series by using, e.g., Lebedev-Laikov quadratures (Kaasalainen et al. 2012). The coefficients $\mathbf{t}_{lm}$ can be used in the following analytical motion-averaged formulae that no longer require convexity, although the relevant $\mathbf{t}_{lm}$ tend to be unstable, as discussed in Sect. 5.

3.1. Rotation speed

Averaging over one rotation, we have

$$\tau(\epsilon,\lambda) = \frac{1}{2\pi}\int_0^{2\pi}\tau(\theta,\varphi)\,\mathrm{d}\varphi = Q\sum_l \mathbf{t}_{l0}\,P_l(\cos\lambda\sin\epsilon). \tag{17}$$

The secular torque $\bar\tau(\epsilon)$ is obtained by averaging over λ:

$$\bar\tau(\epsilon) = \frac{1}{2\pi}\int_0^{2\pi}\tau(\epsilon,\lambda)\,\mathrm{d}\lambda = Q\sum_l \mathbf{t}_{l0}\,I_l(\epsilon), \tag{18}$$

where

$$I_l(\epsilon) = \frac{1}{\pi}\int_0^{\pi}P_l(\cos\lambda\sin\epsilon)\,\mathrm{d}\lambda = \frac{1}{\pi}\int_{-\sin\epsilon}^{\sin\epsilon}\frac{P_l(x)}{\sqrt{\sin^2\epsilon - x^2}}\,\mathrm{d}x. \tag{19}$$

Thus $I_l = 0$ for odd l, and for even l we obtain, using the recursion relations of Legendre polynomials,

$$I_l^{\rm even}(\epsilon) = \frac{2}{\pi}\int_0^{\sin\epsilon} \frac{P_l(x)}{\sqrt{\sin^2\epsilon - x^2}}\,{\rm d}x = C_l\,P_l(\cos\epsilon), \tag{20}$$

where

$$C_l = (-1)^{l/2}\frac{(l-1)!!}{l!!}. \tag{21}$$

Our final expression for $\bar\tau(\epsilon)$ is now simply

$$\bar\tau(\epsilon) = Q\sum_{l\,\rm even} C_l\,\mathbf{t}_{l0}\,P_l(\cos\epsilon), \tag{22}$$

which manifestly vanishes for azimuthally symmetric bodies due to the structure of $\mathbf{T}_{lm}$ discussed above. The torque component that alters the rotation speed of the body is the z-component $\bar\tau_3(\epsilon)$ given by the coefficients $t_{l0}^{(3)}$. If the series is dominated by $t_{20}^{(3)}$, this component essentially vanishes at the roots of $P_2(\cos\epsilon)$, i.e., at $\epsilon \approx 55°$ and $\epsilon \approx 125°$ (cf. NV07).

Shape-related phenomena that are apparent instabilities are obvious from this result. These are actually critical phenomena on a shape close to $\mathbf{t}_{l0} = 0$. Consider, say, a pointed local feature on an otherwise symmetric (e.g., ellipsoidal) body at $\vartheta = \pi/2$, $\psi = \delta > 0$. This feature produces a small $\bar\tau_3 = \tau_f \neq 0$. On the other hand, a similar arrangement symmetric about the x-axis, with the center of the feature at $\psi = -\delta$, yields exactly $\bar\tau_3 = -\tau_f$. Note that δ can be arbitrarily small. One can view this as a minuscule shift of a local feature on the body: practically invisible to the eye, yet the YORP torque changes its sign. This is an analogue of the numerical "boulder shift" phenomenon discussed in Statler (2009).

As an example, we evaluated $\bar\tau_3(\epsilon)$ from Eq. (22) for the convex shapes shown in Fig. 1. In Fig. 2, the various linestyles for different numbers of $T_{l0}^{(3)}$-terms show the rapid convergence of the series. Note that, throughout this paper, we only analyze the shapes of asteroids, so we do not take sizes and masses into account (and set $Q = 1$). Thus the scale of $\bar\tau_3(\epsilon)$ is arbitrary.

3.2. Obliquity

The torque component τ_ϵ that alters the obliquity of the rotation axis is given by the rotational transformation of Eq. (4):

$$\tau_\epsilon = [\mathbf{R}_z(-\varphi')\tau]_1 = \tau_1\cos\varphi' - \tau_2\sin\varphi'. \tag{23}$$

In the same oblique coordinate frame (i.e., the z-axis is aligned with the rotation axis and the x-component of τ is τ_ϵ), the direction of the Sun is given by

$$\omega = \mathbf{R}_y(\epsilon)\begin{pmatrix}\cos\lambda\\ \sin\lambda\\ 0\end{pmatrix} = \begin{pmatrix}\cos\epsilon\cos\lambda\\ \sin\lambda\\ \sin\epsilon\cos\lambda\end{pmatrix} := \begin{pmatrix}\sin\theta\cos\varphi_s\\ \sin\theta\sin\varphi_s\\ \cos\theta\end{pmatrix}. \tag{24}$$

Thus

$$\varphi' = \varphi_s - \varphi, \tag{25}$$

and substituting this to (23) yields

$$\tau_\epsilon = \cos\varphi(\tau_1\cos\varphi_s - \tau_2\sin\varphi_s) + \sin\varphi(\tau_1\sin\varphi_s + \tau_2\cos\varphi_s). \tag{26}$$

From Eq. (24) we find that

$$\cos\varphi_s = \frac{\cos\epsilon\cos\lambda}{\sqrt{1-\sin^2\epsilon\cos^2\lambda}},\quad \sin\varphi_s = \frac{\sin\lambda}{\sqrt{1-\sin^2\epsilon\cos^2\lambda}}, \tag{27}$$

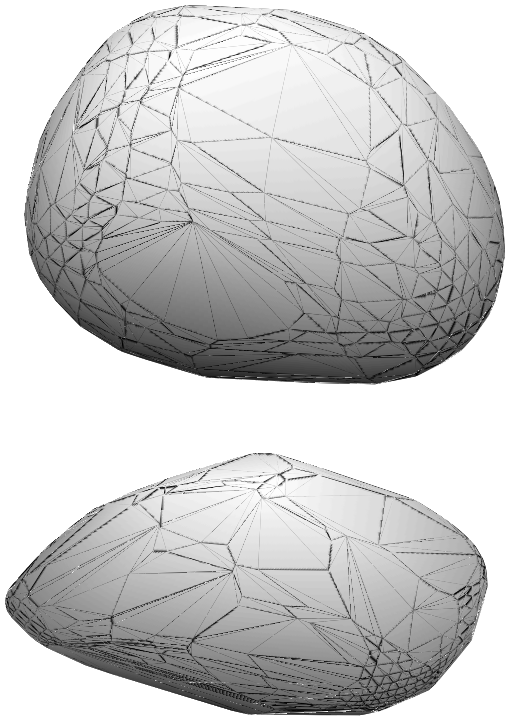

Fig. 1. Convex model shapes of asteroids Psyche (*top*) and Apollo (*bottom*).

so, substituting τ_1 and τ_2 from Eq. (15), using real-valued $\mathbf{t}_{lm}$, and averaging over φ with Eq. (25),

$$\tau_\epsilon(\lambda,\epsilon) = \frac{Q}{2}\sum_l P_l^1(\cos\lambda\sin\epsilon)\left[\cos\varphi_s\left(t_{l1}^{1,\cos} + t_{l1}^{2,\sin}\right)\right. \tag{28}$$
$$\left. +\ \sin\varphi_s\left(t_{l1}^{1,\sin} - t_{l1}^{2,\cos}\right)\right],$$

and averaging over λ, we obtain

$$\bar\tau_\epsilon(\epsilon) = \frac{Q}{2}\sum_l\left(t_{l1}^{1,\cos} + t_{l1}^{2,\sin}\right) I_{l1}(\epsilon) + O, \tag{29}$$

where

$$I_{l1}(\epsilon) = \frac{1}{\pi}\int_0^\pi P_l^1(\cos\lambda\sin\epsilon)\frac{\cos\epsilon\cos\lambda}{\sqrt{1-\sin^2\epsilon\cos^2\lambda}}\,{\rm d}\lambda \tag{30}$$
$$= \frac{1}{\pi}\frac{\cos\epsilon}{\sin\epsilon}\int_{-\sin\epsilon}^{\sin\epsilon}\frac{x\,P_l^1(x)}{\sqrt{1-x^2}}\frac{{\rm d}x}{\sqrt{\sin^2\epsilon - x^2}},$$

so $I_{l1} = 0$ for odd l, and O denotes the $\sin\varphi_s$-terms that vanish because of the parity of P_l^1 and the requirement of an antisymmetric τ_ϵ: $\bar\tau_\epsilon(\epsilon) \equiv -\bar\tau_\epsilon(\pi-\epsilon)$. Computing the integral in the same way as with $\bar\tau(\epsilon)$, we have

$$\bar\tau_\epsilon(\epsilon) = \frac{Q}{2}\sum_{l\,\rm even} C_l\left(t_{l1}^{1,\cos} + t_{l1}^{2,\sin}\right) P_l^1(\cos\epsilon). \tag{31}$$

Owing to the structure of $\mathbf{T}_{lm}$, this vanishes for quadrant-symmetric bodies. In BRV11, Eqs. (29), (C14), and (C15) are alternative versions of Eqs. (15), (22), and (31) here.

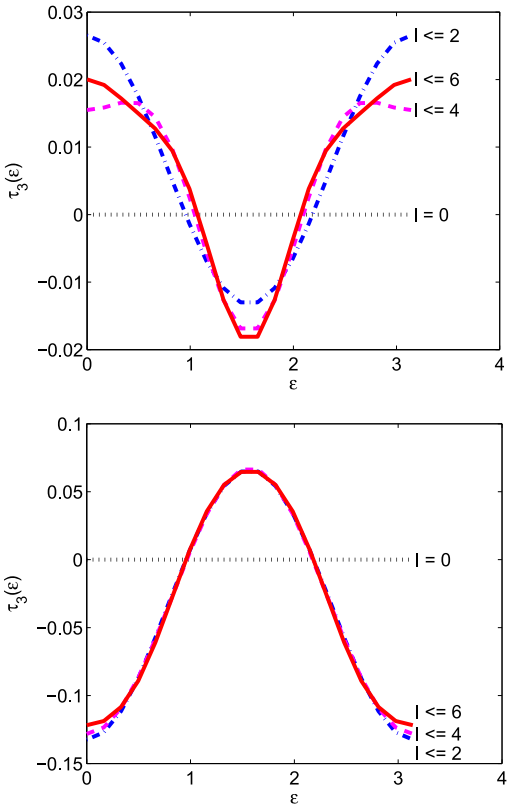

Fig. 2. Secular YORP torque $\bar{\tau}_3(\epsilon)$ for the shapes of Fig. 1 respectively, computed analytically using Eq. (22) for values $l \le 2$, $l \le 4$ and $l \le 6$. Clearly, $l = 2$ is the dominant term, and the series converges rapidly for $l > 2$. The obliquity ϵ is given in radians.

4. Nonzero thermal conductivity

Let us define the insolation factor $\mathrm{ins}(\mu, \xi)$ as $\mathrm{ins}(\mu, \xi) = \mu$ if the point ξ on the surface is illuminated, and otherwise $\mathrm{ins}(\mu, \xi) = 0$. For constant λ and ϵ, $\mathrm{ins}(\mu, \xi)$ is obviously a cyclic function of φ', so it can be expanded as a Fourier series:

$$\begin{aligned}\mathrm{ins}(\mu, \xi) &= \sum_n d_n(\xi, \lambda, \epsilon)\, \mathrm{e}^{\mathrm{i}n\varphi'} \\ &= \sum_{n\ge 0} a_n(\xi, \lambda, \epsilon) \cos n\varphi' + b_n(\xi, \lambda, \epsilon) \sin n\varphi'.\end{aligned} \tag{32}$$

The coefficients $d_n(\xi, \lambda, \epsilon)$ can always be readily computed, analytically for convex bodies and by ray-tracing and FFT for others. In this analysis, however, we do not need their actual values; it is sufficient to know that they exist and fully define $\mathrm{ins}(\mu, \xi)$ (even when it is discontinuous due to shadowing by nonconvexities).

Writing τ by using $\mathrm{ins}(\mu, \xi)$ instead of A_+, we have, for $K = 0$,

$$\begin{aligned}\tau(\lambda, \epsilon, \varphi') &= Q \int_{\mathcal{B}} \mathcal{T}(\xi)\, \mathrm{ins}(\mu, \xi)\, \mathrm{d}S \\ &= Q \sum_n \mathrm{e}^{\mathrm{i}n\varphi'} \int_{\mathcal{B}} \mathcal{T}(\xi) \mathrm{d}_n(\xi, \lambda, \epsilon)\, \mathrm{d}S.\end{aligned} \tag{33}$$

When averaging over φ', the only contributing term is the one with $n = 0$.

For $K \neq 0$, the heat diffusion equation

$$\rho c_p \frac{\partial T}{\partial t} = K \frac{\partial^2 T}{\partial \xi^2}, \tag{34}$$

where ρ is the density of the material, c_p is the specific heat capacity, and the vertical direction ξ is aligned with η, can be solved with suitable physical approximations, boundary conditions, and a periodic ansatz of the form (32) (see NV08). With damping factors Ψ_n and phase lags $\Delta\phi_n$ defined by

$$\Psi_n = \left(1 + 2\Theta_n + 2\Theta_n^2\right)^{-1}, \quad \Delta\phi_n = \mathrm{sgn}(n) \arctan \frac{\Theta_n}{\Theta_n + 1}, \tag{35}$$

where

$$\Theta_n = \frac{\sqrt{\rho c_p}}{4\gamma T_0^3} \sqrt{\frac{1}{2}|n|\, K\, \Omega}, \quad T_0^4 = (1 - A) F_o / \gamma, \tag{36}$$

an approximate solution is

$$T(t)^4\big|_{\xi=0} \simeq (1 - A) F_o \gamma^{-1} \sum_n \Psi_n\, d_n(\eta, \lambda, \epsilon)\, \mathrm{e}^{\mathrm{i}(n\varphi' - \Delta\phi_n)}, \tag{37}$$

so

$$\tau(\lambda, \epsilon, \varphi') = Q \sum_n \Psi_n\, \mathrm{e}^{\mathrm{i}(n\varphi' - \Delta\phi_n)} \int_{\mathcal{B}} \mathcal{T}(\xi) d_n(\xi, \lambda, \epsilon)\, \mathrm{d}S. \tag{38}$$

Again, averaging over φ' retains only the $n = 0$ -contribution, and $\Psi_0 = 1$, $\Delta\phi_0 = 0$, so the result is exactly the same as with $K = 0$. This shows that K has no influence on the rotation-averaged τ.

4.1. Obliquity torque

Using the rotation-transform Eq. (23) for τ_ϵ together with the real-valued Fourier series in (32), we have, from Eqs. (23) and (33),

$$\begin{aligned}\tau_\epsilon(\lambda, \epsilon, \varphi') &= Q \sum_n \Bigg[\cos n\varphi' \cos\varphi' \int_{\mathcal{B}} \mathcal{T}_1(\xi) a_n(\xi, \lambda, \epsilon)\, \mathrm{d}S \\ &\quad - \sin n\varphi' \sin\varphi' \int_{\mathcal{B}} \mathcal{T}_2(\xi) b_n(\xi, \lambda, \epsilon)\, \mathrm{d}S\Bigg] + O,\end{aligned} \tag{39}$$

where O denotes cross-terms that vanish in averaging over φ'. Upon this averaging, we obtain an expression for $K = 0$ ($\Delta\phi_n = 0$) in the Fourier-series representation:

$$\tau_\epsilon(\lambda, \epsilon) = \frac{Q}{2} \int_{\mathcal{B}} \left[\mathcal{T}_1(\xi) a_1(\xi, \lambda, \epsilon) - \mathcal{T}_2(\xi) b_1(\xi, \lambda, \epsilon)\right] \mathrm{d}S, \tag{40}$$

which is thus equivalent to the φ-averaged τ_ϵ of Eq. (28), when assuming constant λ during the averaging.

Next we expand the above expression to $K \neq 0$. With a real-valued Fourier series solution for T^4, averaging $\tau_\epsilon(\lambda, \epsilon, \varphi')$ from Eq. (38) over φ' yields

$$\begin{aligned}\tau_\epsilon(\lambda, \epsilon) &= \frac{Q}{2} \Psi_1 \Bigg[\cos\Delta\phi_1 \int_{\mathcal{B}} (\mathcal{T}_1 a_1 - \mathcal{T}_2 b_1)\, \mathrm{d}S \\ &\quad - \sin\Delta\phi_1 \int_{\mathcal{B}} (\mathcal{T}_2 a_1 + \mathcal{T}_1 b_1)\, \mathrm{d}S\Bigg].\end{aligned} \tag{41}$$

In other words,

$$\tau_\epsilon(\lambda, \epsilon) = \Psi_1 \left[\cos\Delta\phi_1 \tau_\epsilon\left(\tau|_{\Delta\phi=0}\right) + \sin\Delta\phi_1 \tau_\epsilon\left(\tau|_{\Delta\phi=\frac{\pi}{2}}\right)\right], \tag{42}$$

where $\tau_\epsilon(\tau|_{\Delta\phi=\alpha})$ denotes $\tau_\epsilon(\lambda,\epsilon)$ computed with a phase-lagged τ evaluated at $\varphi'-\alpha$ instead of φ'; a phase lag of $\alpha=\pi/2$ changes $a_1\to -b_1$ and $b_1\to a_1$ in the Fourier series. Using $\varphi\to\varphi_s-(\varphi'-\pi/2)$, i.e., $\varphi\to\varphi+\pi/2$, in the fundamental Eq. (15) for τ when averaging over φ and λ, we obtain our final result:

$$\bar\tau_\epsilon(\epsilon) = \frac{Q}{2}\Psi_1 \sum_{l\,\mathrm{even}} C_l\, P_l^1(\cos\epsilon)\left[\cos\Delta\phi_1\left(t_{l1}^{1,\cos}+t_{l1}^{2,\sin}\right)\right. \\ \left. -\sin\Delta\phi_1\left(t_{l1}^{1,\sin}+t_{l1}^{2,\cos}\right)\right]. \quad (43)$$

This means that, at large values of $K\Omega$, the YORP torque $\bar\tau_\epsilon$ acquires a significant component that does not vanish even for ellipsoids, for example.

5. YORP stability

If we consider the coefficients $\mathbf{t}_{lm}$ in the basic torque Eq. (15), we can see that those determining the secular torques are precisely the ones that are usually vanishingly small compared to the momentary torque level. For example, an elongated ellipsoid produces large nonsecular torques, while the secular ones vanish. Thus even a small change in the shape of a body inevitably alters the secular torques with a much larger relative factor. If $\mathbf{t}_{lm}$ are determined numerically from the momentary torques with various illumination directions on S^2, the secular "trash coefficients" are easily affected by their errors, and can even be drowned in the numerical noise. This also helps to understand YORP instability as a fundamental and inevitable fact.

The analytical expressions in Sects. 3 and 4 provide a general set of convenient basis functions for the presentation of numerical YORP computations in a compact form. A robust approach for general shapes, especially those far from convex surfaces, is to compute the secular torques numerically with ray-tracing and averaging, and then fit series in $P_l(\cos\epsilon)$ or $P_l^1(\cos\epsilon)$, l even, to them. This one-dimensional procedure is more stable than the determination of the Laplace series of Eq. (15). One way to compute torques with $K\neq 0$ is to evaluate the cases with emission phase lags of 0 and $\pi/2$ and then form their superposition by modulating with the phase lag $\Delta\phi_1$ to conform with Eq. (43).

Many asteroids are well approximated by convex shapes (up to some level of resolution), so if two shapes differ only slightly, how different are their YORP properties? This question concerns not only the validity regime of an approximation, but the general stability properties of the YORP effect as well.

In this context, it is useful to normalize the YORP torques to be size-invariant, pertaining only to the shape of the object. Since the dimension of the torque integrand is that of volume, we define the dimensionless YORP capacity of a shape as the maximal average torque per unit volume and unit flux pressure (or torque when the volume V and the pressure are scaled to unity):

$$\mathcal{Y} := \frac{\max|\bar\tau_3(\epsilon)|}{QV}, \qquad \mathcal{Y}_\epsilon := \frac{\max|\bar\tau_\epsilon(\epsilon)|}{QV}. \quad (44)$$

The YORP capacity $\mathcal{Y}$ thus describes the amplitude of the $\bar\tau_3(\epsilon)$-curve of the unity-scaled case. For most convex shapes, $\mathcal{Y}$ is fast to evaluate directly from $T_{20}^{(3)}$ with Eq. (22), and $\mathcal{Y}_\epsilon$ similarly with Eqs. (31) or (43). Conversely, given $\mathcal{Y}$ for any shape, one can quickly compute the actual maximal torque by multiplying $\mathcal{Y}$ by QV.

To establish a scale for $\mathcal{Y}$, consider a "propeller" shape consisting of two identical vertically symmetric wedges attached to each other inversion-symmetrically (Fig. 3). The maximal

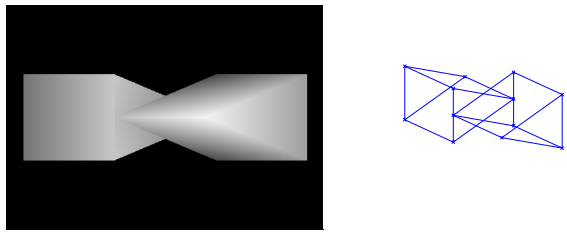

Fig. 3. Wedge-propeller shape strongly prone to the rotational YORP torque, shown as a solid object and as a wire frame.

YORP torque is attained at $\epsilon=0,\pi$. The cross-section of a wedge is an isosceles triangle with height h and width w, and the length of a wedge is l. If we assume a long and thin body with $h/l\ll 1$ (and $w/l\ll 1$), we can neglect the torques from the ends of the wedges and terms of order $(h/l)^2$ and higher, so

$$\hat\tau_3(0)\approx \mathcal{W}\frac{2}{\pi}\int_0^{\pi/2} hl\frac{l}{2}\cos\phi\,d\phi = \mathcal{W}\frac{hl^2}{\pi}, \quad (45)$$

where the efficiency factor $0<\mathcal{W}<1$ of the sharpness of the wedge is

$$\mathcal{W} = 1-\frac{1}{\sqrt{1+(2w/h)^2}}. \quad (46)$$

Thus, with $V=hwl$ and estimating $\mathcal{Y}$ from $|\bar\tau_3(0)|$,

$$\mathcal{Y}\approx\frac{\mathcal{W}}{\pi}\frac{l}{w}\approx 0.1\frac{l}{h} \quad (47)$$

for a flattening ratio $w/h=3/2$ realistic for an asteroid. An infinitely long and thin body yields $\mathcal{Y}\to\infty$, so there is no upper bound for the YORP capacity $\mathcal{Y}$ in principle. For realistic shapes $l/h\sim 1$, we can estimate an upper bound $\mathcal{Y}\lesssim 0.1$.

Any suitably asymmetric, infinitely long and thin shape yields $\mathcal{Y}\to\infty$, so there is no unique answer to the question "what is the most YORP-prone shape", nor is there a unique solution to the problem "what is the most YORP-efficient realistic asteroid shape" since the term "realistic" can be defined in a number of ways. The wedge-propeller shape is more or less the optimal one for YORP-efficiency as it portrays the best "windmill asymmetry" profile.

To study the stability problem, let us estimate an upper bound for the change $|\Delta\mathcal{Y}|$ due to the perturbation $\Delta\boldsymbol{x}$ that affects the characteristic torque function:

$$\mathcal{T}+\Delta\mathcal{T} = (G+\Delta G)(\boldsymbol{x}+\Delta\boldsymbol{x})\times\eta. \quad (48)$$

Consider, e.g., the perturbed unit sphere, so that $\Delta\mathcal{T}\approx\Delta\boldsymbol{x}\times\eta$ (the estimate is essentially similar for all convex bodies, but we choose the sphere for a simple example). Thus

$$\Delta\mathbf{T}_{lm}\approx N_{lm}^{-1}\int_{S^2}\Delta\boldsymbol{x}\times\eta Y_l^{*m}(\eta)\,d\sigma, \quad (49)$$

and we can approximate, with a perturbation $\Delta\boldsymbol{x}=\boldsymbol{f}(\eta)$, where $\boldsymbol{f}(\eta)$ is a Laplace series such that $\boldsymbol{x}+\Delta\boldsymbol{x}$ is convex and $\langle\|\boldsymbol{f}\|\rangle_\eta=\delta$,

$$|\Delta\mathcal{Y}|\sim 2\pi\left|k_2C_2\,N_{20}^{-1}\frac{3}{4\pi}\int_{S^2}(\Delta\boldsymbol{x}\times\eta)_3\,P_2(\vartheta)\,d\sigma\right|\lesssim\delta/10, \quad (50)$$

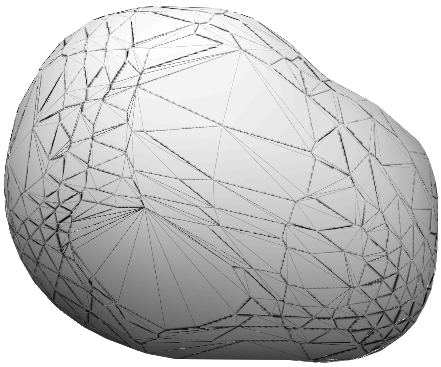

Fig. 4. Nonconvex perturbation based on the model shape of the asteroid Psyche in Fig. 1. Here the value of the YORP capacity $\mathcal{Y}$ of the original body is 1.84×10^{-3}, whereas the perturbed shape has $\mathcal{Y} = 2.06 \times 10^{-3}$. The relative radius perturbation $\langle \Delta r \rangle / \langle r \rangle$ is 0.040.

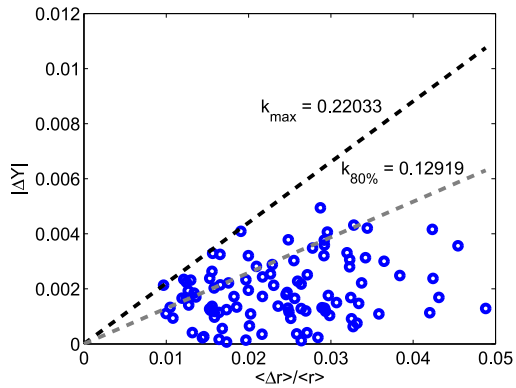

Fig. 5. Plot of $\langle \Delta r \rangle / \langle r \rangle$ versus $|\Delta \mathcal{Y}|$ for different global perturbations of the asteroid Apollo. A bounding line is drawn with the slope $k_{\max} = 0.22$. As the relative perturbation $\langle \Delta r \rangle / \langle r \rangle$ grows, the absolute change in $|\Delta \mathcal{Y}|$ remains bounded and well below the linear upper bound estimate. An alternative bounding line with 80% of the values of $|\Delta \mathcal{Y}|$ remaining under the line is plotted in the same figure, with the slope $k_{80\%} = 0.13$.

where the upper bound for the integral is estimated by applying the Legendre recursion formulae to the products with the components of η and using the orthogonality integral. A perturbation of $\delta = 0.1$ can thus produce a change $\Delta \mathcal{Y}$ roughly a tenth of the wedge propeller's $\mathcal{Y}$ above; in practice, the change is usually smaller than this.

When estimating the change of YORP properties due to perturbations, we only consider the change $|\Delta \mathcal{Y}|$ in amplitude. Indeed, taking the leading Legendre term as in Eq. (50) implies that the changes in the phase (the value of ϵ at the maximum) and the shape of $\bar{\tau}_3(\epsilon)$ are, for small shape perturbations, negligible compared to the amplitude change.

In addition to $\Delta \mathcal{Y}$ by surface fluctuations of a given size (that typically describe the convexity/nonconvexity or resolution difference between two shape models), we can consider a couple of other mechanisms for comparison: i) the "boulder effect" and ii) the effect of scale-invariant surface roughness. If we attach the wedge above to the surface of a sphere of radius r, and fix $h = w = l = \delta r$, we have $\mathcal{Y} \approx 0.01\delta^2$, so adding a boulder of size $r/10$ changes $\mathcal{Y}$ by $\Delta \mathcal{Y} \approx 0.0001$. If we consider a special arrangement of infinitesimal wedges on the surface, all pointing in an ordered way such that $l/h = c$ and $h^2 \to \mathrm{d}\sigma$ and $hl \to c\,\mathrm{d}\sigma$, we obtain an upper bound

$$\frac{3}{4\pi r^3} \frac{1}{\pi} \int_0^{\pi} \int_0^{\pi/2} r \sin\theta \mathcal{W} \cos\phi r^2 \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\phi = \frac{3c\mathcal{W}}{8\pi}, \qquad (51)$$

so a moderate surface roughness $c \approx 0.1$ yields $\Delta \mathcal{Y} \lesssim 0.01$. In principle, one can coat a sphere such that it looks smooth and yet is, paradoxically, quite prone to YORP. The upper bound cannot be reached by a realistic unordered surface structure, but $\Delta \mathcal{Y} \lesssim 0.001$ is plausible. Thus, taking a single large boulder or the unseen small-scale surface structure into account on a small asteroid can change $\mathcal{Y}$ up to $\Delta \mathcal{Y} \lesssim 0.001$.

Given the analytical scale estimates of $\mathcal{Y}$and $\Delta \mathcal{Y}$ above, are there YORP-stable asteroid shapes in reality? We measured the YORP values $\mathcal{Y}$ and their perturbations $\Delta \mathcal{Y}$ for some targets, when the shape perturbations were global nonconvex radius fluctuations Δr of varying magnitude (by low-order Laplace series), and $\mathcal{Y}$ was computed by ray-tracing. A sample perturbation is shown in Fig. 4. For each shape and sequence of perturbations $\Delta r \to \Delta \mathcal{Y}$, one can define an upper bound $k_{\max}$ such that

$$|\Delta \mathcal{Y}| \lesssim k_{\max} \frac{\langle \Delta r \rangle}{\langle r \rangle}. \qquad (52)$$

To estimate this, we plotted $\langle \Delta r \rangle / \langle r \rangle$ vs. $|\Delta \mathcal{Y}|$ for each shape; a sample case is shown in Fig. 5. A bounding line is approximately at $k_{\max} \approx [0.1, 0.2]$, consistent with the upper bound $\delta/10$ estimated above. What is more, the slope $k_{\max}$ decreases as $\langle \Delta r \rangle / \langle r \rangle$ grows ($\Delta \mathcal{Y}$ may even saturate) so the linear upper bound estimate is only a rough approximation for small perturbations. For most shapes, $k_{\max} \sim 0.1$ for a practical estimate. The estimate is not strict, nor can a shape perturbation be properly described by one figure $\langle \Delta r \rangle / \langle r \rangle$ only.

The global perturbation scheme depicts the case of two alternative representations of the same body – for example, a lightcurve-based model and a radar-based one, or two versions of a model based on flyby data. If the model error $\langle \Delta r \rangle / \langle r \rangle$ remains small, we can estimate shapes with $\mathcal{Y} \gtrsim 0.01$ (such as Apollo) to be reasonably YORP-stable against global model errors in the sense of the relative YORP error $|\Delta \mathcal{Y}/\mathcal{Y}|$. Any instability in this error is caused by the inevitable smallness of $\mathcal{Y}$ rather than the size of $|\Delta \mathcal{Y}|$.

Another interesting case is the "boulder effect"; it represents a local model error, or an actual local change of the shape due to some physical perturbation. We made random local changes to bodies and measured the YORP change; a scatter plot similar to that of global perturbations is shown in Fig. 6. In this case, most of the YORP changes are negligible, while some jump to $|\Delta \mathcal{Y}/\mathcal{Y}| \approx 0.1$. As in the analytical estimate above, the relative YORP changes remain well bounded for shapes with $\mathcal{Y} \gtrsim 0.01$ (such as the plotted case of Apollo), while shapes with $\mathcal{Y} \lesssim 0.001$ are no longer YORP-stable against local shape perturbations.

In Table 1, we show sample values of $\mathcal{Y}$ for a number of asteroid shapes (obtained from the DAMIT website). Note again that $\mathcal{Y}$ is size-independent, and the samples represent typical asteroid shapes rather than estimates for the actual bodies, many of which are large and not affected by YORP.

To sum up the analytical and numerical stability estimates above as a rough order-of-magnitude classification scheme, we can define YORP stability classes as

$$\text{stability class} = n = \text{round}\left(\log_{10} \mathcal{Y} + 1.8\right), \qquad (53)$$

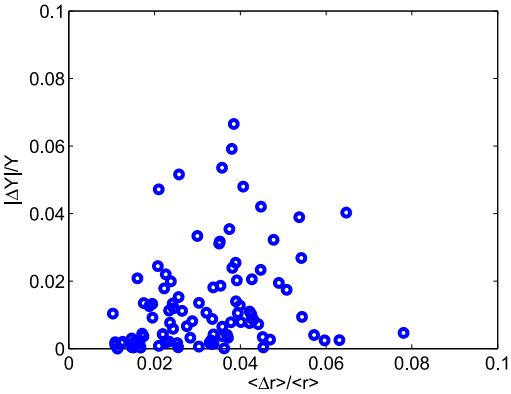

Fig. 6. Plot of $\langle \Delta r \rangle / \langle r \rangle$ versus $|\Delta \mathcal{Y}|/\mathcal{Y}$ for different local perturbations of the asteroid Apollo. Only one facet group of the asteroid is perturbed at a time, and only this section defines $\langle \Delta r \rangle / \langle r \rangle$. With local perturbations less than 10%, the relative change in $\mathcal{Y}$ mostly remains below 10%. Therefore, we can conclude that the asteroid Apollo is locally YORP-stable.

Table 1. List of the YORP capacity $\mathcal{Y}$ of different asteroid shapes.

Asteroid	$\mathcal{Y}$ (10^{-3})
Alexandra	1.60
Apollo	6.61
Daphne	1.55
Eger	8.27
Eunomia	0.72
Euterpe	0.75
Geographos	5.98
Kalliope	1.18
Massalia	5.95
Metis	2.96
Nysa	1.66
Pandora	4.08
Psyche	1.84
Sappho	4.16

Notes. The size is not considered here.

where round(x) gives the integer closest to x, and the constant 1.8 was chosen to set the lower class limit to $\mathcal{Y} = 5 \times 10^{n-3}$. Most asteroid shapes are in the regime $n \leq -1$; some reach $n = 0$, and mostly propeller-type shapes have $n \geq 1$. Depending on the stability class, all the perturbation scenarios above can change $|\Delta \mathcal{Y}/\mathcal{Y}|$ from tens of percent to orders of magnitude for typical asteroid shapes in a quite natural manner. As a ballpark estimate, we could call shapes with $n = 0$ locally YORP-stable and globally semistable, while those with $n = -1$ are locally semistable and globally unstable. All shapes in the classes $n < -1$ are YORP-unstable, and those in $n \geq 1$ YORP-stable.

Most of the shapes in Table 1 are in $n = -1$; only "Massalia", "Apollo", "Geographos", and "Eger" are in $n = 0$. For the last three of these more stable ones, the YORP effect has been observed directly from lightcurves (Kaasalainen et al. 2007; Ďurech et al. 2008, 2012). The stability estimate is consistent with the good correspondence between the computed and observed YORP acceleration for Apollo and Geographos. For Eger, the correspondence is not as good. The value of $\mathcal{Y}$ thus indicates that this is more due to its unknown sizable nonconvex features than the smaller-scale irregularities (the cited value is based on the adopted nonconvex shape solution; the convex one has an even higher $\mathcal{Y} = 0.012$). Indeed, Eger is notable for being one of the few asteroids for which lightcurves indicate strong nonconvexities (Ďurech & Kaasalainen 2003). Some main features can be sketched from photometric data (Ďurech et al. 2012), but the solution is not unique, which may cause the discrepancy between the expected and real YORP strengths.

As in the analytical estimates, the amplitude change $|\Delta \mathcal{Y}/\mathcal{Y}|$ dominated our numerical perturbation cases. This means that, for most of the $\bar{\tau}_3(\epsilon)$-curve, $\mathcal{Y}$ determines the stability class. However, when ϵ is close to the values at which $\bar{\tau}_3(\epsilon)$ vanishes (i.e., usually around $\epsilon \approx 55°$ and $\epsilon \approx 125°$), the changes in the phase and shape of the curve induced by the shape perturbation are significant for the relative change in $\bar{\tau}_3(\epsilon)$. Near these obliquities, all shapes are YORP-unstable.

6. Conclusions and discussion

We have presented a concise formalism for analyzing YORP-related quantities. Using the characteristic torque function of the body, defined by Eq. (8), one can readily write the general Eq. (15) for the YORP torque as well as its main secular components of interest: the one changing the rate of rotation and the one altering the obliquity. The motion-averaged results can be given for both zero (Eqs. (22), (31)) and nonzero (Eq. (43)) thermal inertia. Since the averaging integrals are analytical regardless of the shape, the formulae also provide a convenient set of basis functions for representing YORP computations of nonconvex shapes.

The YORP effect as such is stable in principle; i.e., a small change in the shape leads to a small change $\Delta \mathcal{Y}$ in the YORP capacity $\mathcal{Y}$. What causes the unpredictability is the relative change $\Delta \mathcal{Y}/\mathcal{Y}$ because most asteroids tend to have $\mathcal{Y}$ close to zero. Here we have derived a number of analytical and numerical estimates on YORP stability, and, based on these, surmise that most asteroids tend to be YORP-unstable, though some can reach a stable class. The stability class is determined by the YORP capacity $\mathcal{Y}$: the larger the capacity, the better the relative YORP stability. For convex polyhedra, a simple approximation of $\mathcal{Y}$ is obtained from the P_2-coefficient of $\bar{\tau}_3(\epsilon)$:

$$\mathcal{Y} \approx \frac{5}{64V} \Big| \sum_i A_i \, (\boldsymbol{x}_i \times \eta_i)_3 \left(3 \cos^2 \vartheta_i - 1 \right) \Big|, \tag{54}$$

where A_i, $\boldsymbol{x}_i$, and η_i are the area, centroid, and unit normal of a facet, and ϑ_i is the polar angle of the normal (see Appendix). This approximation can also be used directly for slightly nonconvex shapes to obtain a first estimate; for these, the sum over facets is almost similar to a convex case.

Here we have simplified the YORP business to the extreme at the expense of accuracy, but Eq. (54) provides some equational economy: with it (and its variants for other torques), we can approximate $\bar{\tau}_3(\epsilon)$ and $\mathcal{Y}$ and estimate the YORP stability of a body. For more strongly nonconvex shapes, one can construct the best-fit convex shape by, e.g., scaling the convex hull suitably (semianalytical generalizations of NV07 via a perturbation of this shape are also relatively straightforward, but they lose the conceptual simplicity with no certainty of better accuracy). If the fluctuations around this surface do not exceed, say, a tenth part of the radius, the analytical result is reasonably reliable if $\mathcal{Y} \gtrsim 0.01$. For $\mathcal{Y} \sim 0.001$, the accuracy of any YORP estimate (not just the types discussed here) is uncertain due to the instability. For $\mathcal{Y} \lesssim 0.0001$, all YORP estimates are likely to

be invalid, and only upper bounds can be given. The same applies to YORP estimates made near the obliquity at which the averaged net torque vanishes, regardless of $\mathcal{Y}$.

An unstable class does not mean that all shape perturbations induce a large change in $\mathcal{Y}$, but that suitable ones can do so. It would be interesting to explore the YORP stability properties in "shape space" by evaluating $\mathcal{Y}$ and $\mathcal{Y}_\epsilon$ for a large number of both convex and nonconvex shapes and their perturbations by craters and valleys. This would shed more light on the role of the YORP effect in the evolutionary history of asteroids. For example, a small asteroid with a low YORP capacity is likely to undergo a more stochastic YORP-affected spin evolution than one with a higher value of $\mathcal{Y}$, since even small alterations of the shape may be significant for the YORP effect.

Acknowledgements. This work was supported by the Academy of Finland project "Modelling and applications of stochastic and regular surfaces in inverse problems". We thank David Vokrouhlický for valuable comments that improved the manuscript.

Appendix A: Polyhedral representations

The torque over a planar facet of area A_i is given by the torque density at the centroid of the facet multiplied by the facet area, as is easily verified by taking the integral (13) on the facet alone in a coordinate system transformed to the facet plane. Thus, e.g., the YORP torque of a convex polyhedron at $K = 0$ is

$$\tau = Q \sum_i \frac{A_i}{3} \mu_i \sum_{j=1}^{3} \boldsymbol{x}^{(ij)} \times \eta_i, \tag{A.1}$$

where $\boldsymbol{x}^{(ij)}$ are the vertices of the triangle facets, and their areas and outward unit normals are A_i and η_i.

Consider next

$$\mathbf{T}_{lm} = N_{lm}^{-1} \int_{S^2} G(\eta)(\boldsymbol{x} \times \eta) Y_l^{*m}(\eta)\, \mathrm{d}\sigma, \tag{A.2}$$

where

$$N_{lm} = \frac{4\pi}{2l+1} \frac{(l+m)!}{(l-m)!} \tag{A.3}$$

for the unnormalized $Y_l^m(\vartheta, \psi) = P_l^m(\cos\vartheta)\mathrm{e}^{\mathrm{i}m\psi}$. In the real-valued version of the spherical harmonics series, the coefficients for the sine and cosine terms with $m > 0$ are twice the ones in the complex-valued case. The torque from Eq. (A.1) can be interpreted as the Dirac-delta limit of the function-based form of Eq. (13), so we obtain an exact result for the coefficients $\mathbf{T}_{lm}$ of a convex polyhedron by applying the same limit to Eq. (A.2):

$$\mathbf{T}_{lm} = N_{lm}^{-1} \sum_i A_i (\boldsymbol{x}_i \times \eta_i) Y_l^{*m}(\eta_i), \tag{A.4}$$

where $\boldsymbol{x}_i$ denotes the centroid of the facet i. If $G(\eta)$ and $\boldsymbol{x}(\eta)$ are defined with spherical harmonics series, τ or $\mathbf{T}_{lm}$ can be computed analytically as mentioned earlier, or with quadratures, especially Lebedev-Laikov on S^2 (Kaasalainen et al. 2012).

The change in $\mathcal{T}$ due to shifting the origin to some $\boldsymbol{x}_0$; i.e., $\boldsymbol{x} \to \boldsymbol{x} - \boldsymbol{x}_0$, is easily computed by writing $\mathcal{T} \to \mathcal{T} - \Delta\mathcal{T}$ with

$$\Delta\mathcal{T}(\eta) = G(\eta)\, \boldsymbol{x}_0 \times \eta. \tag{A.5}$$

Thus the coefficients $\mathbf{T}_{lm}$ change as $\mathbf{T}_{lm} \to \mathbf{T}_{lm} - \Delta\mathbf{T}_{lm}$, with

$$\begin{aligned} \Delta\mathbf{T}_{lm} &= N_{lm}^{-1} \int_{S^2} G(\eta)\, (\boldsymbol{x}_0 \times \eta) Y_l^{*m}(\eta)\, \mathrm{d}\sigma \\ &= N_{lm}^{-1} \sum_i A_i (\boldsymbol{x}_0 \times \eta_i) Y_l^{*m}(\eta_i), \end{aligned} \tag{A.6}$$

which, if $G(\eta)$ is written as a spherical harmonics series, is simple to evaluate by writing the products of Y_l^m and the components of η with the Legendre recursion relations and using the orthogonality integral.

References

Bottke, W. F., Vokrouhlický, D., Broz, M., et al. 2001, Science, 294, 1693
Breiter, S., & Michalska, H. 2008, MNRAS, 388, 927
Breiter, S., Rozek, A., & Vokrouhlický, D. 2011, MNRAS, 417, 2478 (BRV11)
Chesley, S. R., Ostro, S. J., Vokrouhlický, D., et al. 2003, Science, 302, 1739
Ďurech, J., & Kaasalainen, M. 2003, A&A 404, 709
Ďurech, J., Vokrouhlický, D., Kaasalainen, M., et al. 2008, A&A, 489, L25
Ďurech, J., Vokrouhlický, D., Baransky, A. R., et al. 2012, A&A, 547, A10
Kaasalainen, M., & Lamberg, L. 2006, Inverse Problems, 22, 749
Kaasalainen, M., Lamberg, L., Lumme, K., & Bowell, E. 1992, A&A, 259, 318
Kaasalainen, M., Ďurech, J., Warner, B., et al. 2007, Nature, 446, 420
Kaasalainen, M., Lu, X., & Vänttinen, A. 2012, A&A, 539, A96
Lowry, S. C., Fitzsimmons, A., Pravec, P., et al. 2007, Science, 316, 272
Nesvorný, D., & Vokrouhlický, D. 2007, ApJ, 134, 1750 (NV07)
Nesvorný, D., & Vokrouhlický, D. 2008a, ApJ, 136, 291 (NV08)
Nesvorný, D., & Vokrouhlický, D. 2008b, A&A, 480, 1
Rozitis, B., & Green, S. 2012, MNRAS, 423, 367
Statler, T. 2009, Icarus, 202, 502
Vokrouhlický, D., Nesvorný, D., & Bottke, W. 2003, Nature, 425, 147
Vokrouhlický, D., Brož, M., Bottke, W. F., et al. 2006, Icarus, 182, 118

Publication II

Nortunen, H., Kaasalainen, M., Ďurech, J., Cibulková, H., Alí-Lagoa, V., and Hanuš, J., "Shape and spin distributions of asteroid populations from brightness variation estimates and large databases," *Astronomy and Astrophysics*.

A&A 601, A139 (2017)
DOI: 10.1051/0004-6361/201629850

Astronomy & Astrophysics

Shape and spin distributions of asteroid populations from brightness variation estimates and large databases

H. Nortunen[1], M. Kaasalainen[1], J. Ďurech[2], H. Cibulková[2], V. Ali-Lagoa[3,4], and J. Hanuš[2]

[1] Tampere University of Technology, Department of Mathematics, PO Box 553, 33101 Tampere, Finland
e-mail: hari.nortunen@tut.fi
[2] Charles University, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovickách 2, 18000 Praha 8, Czech Republic
[3] Observatoire de la Côte d'Azur, Boulevard de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[4] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany

Received 5 October 2016 / Accepted 1 March 2017

ABSTRACT

Context. Many databases on asteroid brightnesses (e.g. ALCDEF, WISE) are potential sources for extensive asteroid shape and spin modelling. Individual lightcurve inversion models require several apparitions and hundreds of data points per target. However, we can analyse the coarse shape and spin distributions over populations of at least thousands of targets even if there are only a few points and one apparition per asteroid. This is done by examining the distribution of the brightness variations observed within the chosen population.

Aims. Brightness variation has been proposed as a population-scale rather than individual-target observable in two studies so far. We aim to examine this approach rigorously to establish its theoretical validity, degree of ill-posedness, and practical applicability.

Methods. We model the observed brightness variation of a target population by considering its cumulative distribution function (CDF) caused by the joint distribution function of two fundamental shape and spin indicators. These are the shape elongation and the spin latitude of a simple ellipsoidal model. The main advantage of the model is that we can derive analytical basis functions that yield the observed CDF as a function of the shape and spin distribution. The inverse problem can be treated linearly. Even though the inaccuracy of the model is considerable, databases of thousands of targets should yield some information on the distribution. We employ numerical simulations to establish this and analyse photometric databases that provide sufficiently large numbers of data points for reliable brightness variation estimates.

Results. We establish the theoretical soundness and the typical accuracy limits of the approach both analytically and numerically. We propose a robust brightness variation observable η based on at least five brightness points per target. We also discuss the weaker reliability and information content of the case of only two points per object. Using simulations, we derive a practical estimate of the model distribution in the (shape, spin)-plane. We show that databases such as Wide-field Infrared Survey Explorer (WISE) yield coarse but robust estimates of this distribution, and as an example compare various asteroid families with each other.

Key words. methods: statistical – methods: numerical – techniques: photometric – minor planets, asteroids: general – methods: analytical

1. Introduction

Most of the current roughly one thousand asteroid shape and spin models, such as the ones given in the Database of Asteroid Models from Inversion Techniques (DAMIT[1]), are based on photometry (Kaasalainen & Lamberg 2006; Ďurech et al. 2015). Databases from large sky surveys and the inversion methods of sparse lightcurves (Kaasalainen 2004; Ďurech et al. 2009, 2016a) will greatly expand the list of models of individual asteroids. However, the databases also contain measurements that are not sufficient for individual models, but nevertheless can be expected to provide information on the statistical shape and spin distributions of the observed asteroid populations. Such measurements are, for example, brightness sequences ranging from a few points to full lightcurves. These can be transformed into statistical data by examining the population-level distribution of the brightness variation within each observed sequence.

The variation among each target's brightnesses, sampled over a wide range of rotation phases, can mostly be attributed to the shape elongation and the sub-Earth aspect angle of the object. The more detailed shape of the body, and especially its irregularity, is another important factor, but this cannot be included in a population-level model due to its complexity. The elongation and aspect have simply describable effects on the brightness variation, with monotonous dependencies. The detailed illumination and viewing geometry also have a somewhat complicated effect. Fortunately, if we include these factors as a part of the modelling error by simply using opposition geometry in the model, they will not make a large contribution to the total error budget as we will discuss below.

A realization of the statistical approach was presented by Szabo & Kiss (2008). They included over 10^4 pieces of pairwise brightness differences and, while their study did not contain analytical or numerical inspection of the generic inverse problem, they concluded that a statistical analysis is possible. The asteroid populations were characterized by shape elongation

[1] http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php

distributions. A similar type of observable and method was used by McNeill et al. (2016). We aim to establish the usefulness of the statistical approach by investigating the inverse problem both analytically and numerically by simulations, including the role of the insufficient model and other assumptions that do not necessarily hold in practice. We also seek to define a good observable of the brightness variation such that its information content is as high as possible.

We generate cumulative distribution functions (CDFs) of the brightness variation levels observed within large asteroid populations, aiming to study what the CDF reveals about the properties of the population. We choose the CDF since it is the most direct, well-defined, and stable data product describing the distribution statistics of a one-dimensional observable. Morever, the analytical study of the inverse problem requires the CDF integral of the model in the first place. To keep our model CDF simple and solvable, we choose to utilize as few parameters as possible, namely the shape elongation p and the spin β. We define p as the ratio of the equatorial widths of the asteroid, and β is the ecliptic polar angle of the spin axis. The main principle is to derive analytical basis functions that describe the contribution of the proportion of targets in a given β and p-bin to the observed CDF. These functions allow both the inspection of the information content of the data and the use of robust inversion methods. In particular, we show that it is possible to obtain information about the β distribution in addition to p.

In addition to the thorough analysis of the inverse problem, one of our main goals is to introduce an especially useful observable, η, that is a measure of the variation of the squared intensities of a sequence. The estimate η can be employed in a variety of contexts. Cibulková et al. (2016) used η to investigate brightness data that were not sufficient for sparse lightcurve inversion but suitable for creating a number of most probable simple asteroid models. These were used especially to demonstrate the slight periodic anisotropy of the distribution of rotation longitudes.

As the shape and spin distributions of asteroid populations are complex to interpret by themselves, we aim to introduce a tool for comparing the distributions of different populations. We do not make astronomical interpretations of the populations or their differences, but our objective is to show that the CDF-based method is a useful tool for the statistical investigation of populations.

We use asteroid databases that provide a number (usually at least five) of points for effectively one rotation: that is, obtained essentially randomly within a few nights such that the aspect angles of the Earth and the Sun are effectively constant. This allows the use of analytical basis functions for the CDFs as well as a rigorous study of the inverse problem. Other scenarios can be used as well, but these require additional assumptions and/or purely numerical treatment, further increasing the model noise.

In Sect. 2 we formulate the observables and the forward problem of the derivation of a CDF from the population model. In Sect. 3 we discuss the solution methods of the inverse problem and prove its fundamental uniqueness and stability properties. In order to verify the applicability of our method and obtain information on the level of error, we perform realistic simulations in Sect. 4 to assess the information content in practice. In Sect. 5 we use observations from databases to analyse asteroid families. We do not consider observational biases here: we simply take the available data at face value and analyse them as such. Additionally, we introduce a tool for a statistical comparison of distinct families. Conclusions are presented in Sect. 6, and mathematical details are given in Appendices A to C.

2. Observables and forward problem

In this section, we define our model and formulate the forward problem. We introduce our main observable, denoted as η, based on the variation of squared brightness intensities, and show why it is useful both analytically and by its information content. We also briefly consider the applicability of measurements with only two observed points. Mathematical details are given in Appendices A and B.

Our model shape is the triaxial ellipsoid, since it has a particularly simple analytical expression for the area of its projection in any given viewing direction (Connelly & Ostro 1984). In this paper, we use the terms brightness and projection area interchangeably, because they are physically almost the same (up to a scaling factor) for dark targets when the viewing and illumination directions coincide (Kaasalainen & Lamberg 2006). We further simplify the model (semiaxes a, b, c) with $b = c = 1$ and use $p := b/a$ for describing the shape elongation (the smaller the p, the more elongated the body). This is a coarse shape approximation for individual targets of general shape, but even if our model is actually not very realistic in practice, it should portray some coarse-scale population tendencies correctly when we have many observations. Indeed, as we will show by simulations, it suffices to have a model that represents the effects of shape elongation and spin direction in a roughly correct manner.

2.1. Amplitude A and its CDF C(A)

Let the polar aspect angle of the viewer be given by θ: $\cos\theta = \boldsymbol{v}\cdot\boldsymbol{e}$, where $\boldsymbol{v}$ is the spin direction (given by the polar coordinates (β, λ) in the inertial frame) and $\boldsymbol{e}$ the line of sight (unit vectors). Due to model symmetry, we only need to consider the interval $0 \le \theta \le \pi/2$. With ϕ for the longitudinal angle in a coordinate frame fixed to the ellipsoid, the area I of the ellipsoid's projection in the direction $\boldsymbol{e}$ is (Connelly & Ostro 1984)

$$I = \pi abc\sqrt{\frac{\sin^2\theta\cos^2\phi}{a^2} + \frac{\sin^2\theta\sin^2\phi}{b^2} + \frac{\cos^2\theta}{c^2}}.$$

In terms of our model definitions, the brightness L scaled against the maximal possible value πa is

$$\begin{aligned} L &= \sqrt{p^2\sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta} \\ &= \sqrt{1 + (p^2-1)\sin^2\theta\cos^2\phi}. \end{aligned} \tag{1}$$

The statistical observable can be anything that describes the variation of the brightness as the target rotates (at a fixed θ). A simple version is the peak-to-peak amplitude; here we consider the ratio $A = L_{\min}/L_{\max} = L|_{\phi=0}/L|_{\phi=\pi/2}$ (i.e. an "inverse amplitude": the smaller the A, the larger the variation). Thus we have chosen the convenient $0 < p \le 1$ and $0 < A \le 1$ (rather than either of these extending to infinity). We would like to note that the amplitude A is based on intensity; we do not use magnitudes anywhere. The assumption is that all objects rotate about an axis (the ellipsoid's c-axis), which produces the observed projections random in ϕ. At first, we consider the randomness of θ to be due to the uniform distribution of rotation axis directions on the unit sphere S^2; later, we take the θ-distribution to be caused by a shifting viewing position.

The amplitude A is given by

$$A = \sqrt{\cos^2\theta + p^2\sin^2\theta} = \sqrt{1 + (p^2-1)\sin^2\theta}. \tag{2}$$

Using the amplitude, we can derive analytical basis functions, the linear combination of which yields the CDF $C(A)$ of a population with a given distribution of p (and β); see Appendix A for details.

In an approximation consistent with the coarseness of the model, it is practical to divide the population under study into a moderate number n of bins in each of which all members have the same p (and β). Then, if we have only p-bins and isotropic θ, the CDF of the values of A observed in the model population is

$$C(A) = \sum_{i=1}^{n} w_i F_i(A), \tag{3}$$

where the basis functions $F_i(A)$ are, from Eq. (A.2),

$$F_i(A) = \begin{cases} 0, & A \le p_i \\ \sqrt{\frac{A^2-p_i^2}{1-p_i^2}}, & A > p_i. \end{cases} \tag{4}$$

The range of the monotonously increasing F_i is $[0, 1]$, and $F_i = 1$ at $A = 1$ (Fig. 1). The occupation numbers of the bins are given by w_i.

Let us now include the β-distribution[2] by assuming that there is a concentration of viewing geometries towards the ecliptic plane (see Appendix A). If we assume a (p_i,β_j)-grid, $i = 1, \ldots, l$ and $j = 1, \ldots, m$, then we have $n = lm$ bins in the grid. We can write the CDF $C(A)$ as

$$C(A) = \sum_{ij} w_{ij} F_{ij}(A), \tag{5}$$

where, from Eq. (A.4), the monotonously increasing basis functions $F_{ij}(A)$ with the range $[0, \pi/2]$ are,

$$F_{ij}(A) = \begin{cases} 0, & A \le p_i \\ \frac{\pi}{2} - \arccos \frac{\sqrt{A^2-p_i^2}}{\sin\beta_j \sqrt{1-p_i^2}}, & p_i < A < \mathcal{F}(p_i,\beta_j) \\ \frac{\pi}{2}, & A \ge \mathcal{F}(p_i,\beta_j), \end{cases} \tag{6}$$

where $\mathcal{F}(p_i,\beta_j) = \sqrt{\sin^2\beta_j + p_i^2\cos^2\beta_j}$. The $F_{ij}(A)$ are sigmoidal functions (Fig. 2), approaching the step function when $p_i \to 1$ (step at $A = 1$) or $\beta_j \to 0$ (step at $A = p_i$). Because of our choice of scale of p and A, parts of the F_{ij} tend to pack together at the low end of A, making them less well distinguishable than those with the slope in the higher end of A, but on the other hand, p-values less than 0.4 are not likely for real celestial bodies.

The occupation numbers w_{ij} are assigned to each bin. It should be noted that occupation levels proportional to $\sin\beta$ mean a uniform density on the direction sphere: that is, a constant $f(\beta)$. For applications, we adopt the convention of reporting the actual (relative) target numbers w_{ij} for a given β-slot (absorbing the factor $\sin\beta$), and we plot these as the density functions (DF; number densities in β rather than on the sphere) in the following sections.

[2] We note that our β is measured from the pole: $0 \le \beta \le \pi$. The ellipsoidal model, however, folds here all solutions of β into the interval $0 \le \beta \le \pi/2$; that is, the model cannot distinguish between pole latitudes above and below the ecliptic plane.

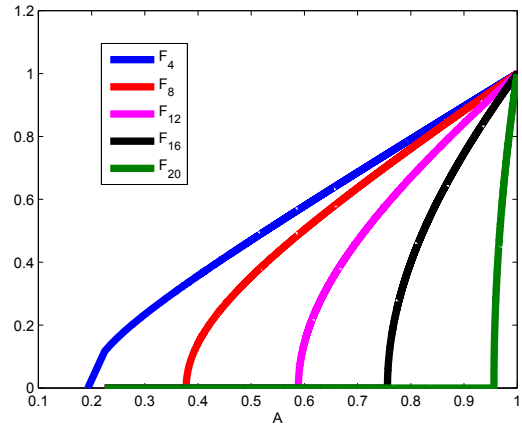

Fig. 1. Sample basis functions F_i on a set of bins p_i, where $i = 1, \ldots, 20$.

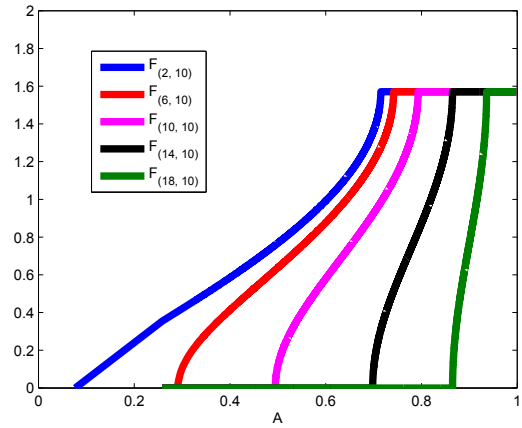

Fig. 2. Sample basis functions F_{ij} on a set of bins (p_i,β_j), where $i = 1, \ldots, 20$ and $j = 1, \ldots, 19$. The shape of the basis functions shows that they are linearly independent.

2.2. Brightness variation η

If the amplitude cannot be measured directly, a practical observable is the brightness variation around some mean value, requiring fewer points. Using intensity squared, L^2, to get rid of the square root in integrands, we obtain from Eq. (1) a simple average quantity over model rotation at a constant θ:

$$\begin{aligned} \langle L^2 \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \sin^2\theta(p^2-1)\cos^2\phi\right) \mathrm{d}\phi \\ &= 1 + \frac{1}{2}\sin^2\theta(p^2-1). \end{aligned}$$

Now, a measure of variation[3] for L^2 over a rotation is

$$\begin{aligned} \Delta(L^2) &= \sqrt{\langle (L^2 - \langle L^2 \rangle)^2 \rangle} = \sqrt{\langle [\sin^2\theta(p^2-1)(\cos^2\phi - 1/2)]^2 \rangle} \\ &= \sin^2\theta(1-p^2)\left[\frac{1}{2\pi}\int_0^{2\pi} (\cos^4\phi - \cos^2\phi)\,\mathrm{d}\phi + \frac{1}{4}\right]^{1/2} \\ &= \sin^2\theta(1-p^2)/\sqrt{8}, \end{aligned}$$

[3] Other definitions could be used as well, but this form leads to simple closed-form formulae.

and normalizing this with $\langle L^2\rangle$ yields

$$\eta(\theta,p) := \Delta(L^2)/\langle L^2\rangle = \sqrt{\left\langle\left(\frac{L^2}{\langle L^2\rangle}-1\right)^2\right\rangle} = \frac{1}{2\sqrt{2}}\left[\frac{1}{\sin^2\theta(1-p^2)}-\frac{1}{2}\right]^{-1}. \quad (7)$$

We note that $0 \le \eta \le 1/\sqrt{2}$. Thus, by Eq. (2), our brightness variation η is directly related to the amplitude A:

$$\eta = \frac{1}{\sqrt{8}}\left(\frac{1}{1-A^2}-\frac{1}{2}\right)^{-1}, \quad A = \sqrt{1-\left(\frac{1}{\sqrt{8}\eta}+\frac{1}{2}\right)^{-1}}. \quad (8)$$

This is a particular advantage of the biaxial model: we can use all available estimates of A (available for dense lightcurves) and η together to form a $C(A)$. One can also directly compute a $C_\eta(\eta)$ with a procedure similar to that of Appendix A, resulting in similar types of integrals, but we choose the A-based formulation as it is more intuitive and leads to simpler equations. The end result is naturally the same in both cases. For the triaxial ellipsoid, a similar simple conversion between A and η is not possible since η would depend on θ (and b) in addition to A (Cibulková et al. 2016).

The condition $0 \le \eta \le 1/\sqrt{2}$ from Eq. (7) may be violated at some measurements of $L(\phi,\theta,p)$ when the parameter p is low ($\lesssim$0.4). The maximal theoretical value of η for all lightcurve shapes (not just those from ellipsoids) approaches one, given by a boxcar-shaped lightcurve with half of the values at a constant level and half approaching zero. Real lightcurves have lower values of η because the lightcurve is smoother than the step-function type. If $\eta > 1/\sqrt{2}$ (this may happen due to an irregular shape, outliers, and/or particular spacing of the sample points), it follows that the amplitude A becomes purely imaginary according to Eq. (8). For computational purposes, we have omitted complex amplitudes in our study (these are rarely encountered).

2.3. Two-point brightness variation

The accuracy of the η estimate depends on the number of data points (and their coverage of the rotational phase) used to approximate $\Delta L^2/\langle L^2\rangle$. To analyse the information content of the minimal case of two points per rotation, we briefly consider simple pairwise brightness differences. For a group of N points for one target, the number of such values is $N(N-1)/2$, ordered such that the difference $0 < q \le 1$ is $q = L_{\rm dimmer}/L_{\rm brighter}$. We do not need to have more than one such pair for one target, so one object does not have to cover the rotational phases well. This is the observable used in Szabo & Kiss (2008). We examine its properties from the inversion point of view in Appendix B. Since they turn out to be considerably inferior to those of η (above all, no information can be obtained on the distribution of β), we do not consider the two-point data further in the main text or database analysis.

McNeill et al. (2016) used a similar type of observable, with the two points connected by a short time interval (effectively yielding the slope of a lightcurve). This problem is even more complicated as it necessarily introduces the rotation period of the target into the forward model, with overlapping effects of p, spin, and period distributions (see Appendix B). Thus, in practice, this observable necessitates the heavy use of a priori assumptions in the inverse problem. Again, this is outside our aim of minimal use of parameters and prior functions, so we do not consider the slope version of two-point data further.

3. Inverse problem

In this section, we consider the fundamental properties of the inverse problem version of the forward model above before moving to realistic shapes and numerical results in the following sections. In particular, we present and prove a uniqueness result that shows how the distributions of both p and β can be uniquely obtained from the CDF $C(A)$. That is, we show why η-data contain unambiguous information on $f(p,\beta)$. This may seem counterintuitive at first glance, since the effects of p and β are certainly mixed for a single observation of η (i.e. a lightcurve). The point is that, under the ecliptic-plane assumption of Appendix A, the distribution of η in a large population separates the effects from each other.

For the p-only case of isotropic θ, the inverse problem can be cast linearly in matrix form. From Eq. (3), we write

$$Mw = C, \quad (9)$$

where $C \in \mathbb{R}^k, w \in \mathbb{R}^n$, $M_{ji} = F_i(A_j)$, and $F_i(A_j)$ are given by Eq. (4). The k observed values of A (derived from η) are sorted in ascending order, and the vector C contains the observed CDF: each element $C_j = j/k$ is the value of $C(A_j)$. In Appendix C, we discuss the analytical stability results of the distribution function $f(p)$ obtained from $C(A)$ (i.e. η-scatter data). In particular, we show that the inverse problem is not strongly ill-posed: the errors in the details of the observed CDF do not amplify fast in the error of the recovered $f(p)$.

In the inverse problem of full $f(p,\beta)$, we can write Eq. (5) in the form of Eq. (9) as well; now $C \in \mathbb{R}^k$, $w \in \mathbb{R}^n$ and M is a $k \times n$-matrix ($n = lm$),

$$M = \begin{pmatrix} F_{11}(A) & \dots & F_{1m}(A) & \dots & F_{l1}(A) & \dots & F_{lm}(A)\end{pmatrix},$$

and the occupation numbers w_{ij} are given in w with indexing similar to M above.

Uniqueness result. Perhaps surprisingly, the distribution of both p and β is unambiguously recoverable; that is, the bin model coefficients w_{ij} are uniquely determined by the $C(A)$. To show this, we notice that the pairs of end points (i.e. the values of A between which F_{ij} changes: A_- at $F_{ij} = 0$ and A_+ at $F_{ij} = \pi/2$), are unique for each F_{ij}. Any combination of F_{ij} will start to deviate from zero at the lowest A_- of the set, and stop changing at the highest A_+ of the set. Thus both end points of an F_{ij} cannot be matched by a superposition of other F_{rs}, so the F_{ij} are linearly independent (see Fig. 2 for illustration). Since the model $C(A)$ is a linear combination of the F_{ij}, the w_{ij} are unique for the observed $C(A)$. As w_{ij} are the occupation numbers of each (p_i,β_j) bin, this proves that the full p and β distribution is uniquely obtained for the ecliptic-orbit model.

If we want to use regularization to smooth the solutions for either p or β, we may apply, for example, the following $(n-1)\times n$ regularization matrix in the p-only case:

$$(R_p)_{ij} = \begin{cases} -1/(p_{i+1}-p_i), & i = j \\ 1/(p_{i+1}-p_i), & j = i+1 \\ 0, & \text{elsewhere}\end{cases}$$

and its generalization for the (p,β)-grid, as well as similarly R_β with β. These approximate the gradients at each w_{ij} in the p- and β-directions only; one can construct more general matrices, but we found these to suffice for our problem. The occupation numbers can be obtained as a solution to an optimization problem:

$$\hat{w} = \arg\min_w\left(\|C-Mw\|^2 + \delta_p\|R_p w\|^2 + \delta_\beta\|R_\beta w\|^2\right), \quad w \in \mathbb{R}^n_+. \quad (10)$$

To obtain the solution $\hat{w}$, we create an extended matrix $\tilde{M}$:

$$\tilde{M} = \begin{pmatrix} M \\ \sqrt{\delta}_p R_p \\ \sqrt{\delta}_\beta R_\beta \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} C \\ 0_{(l-1)m} \\ 0_{l(m-1)} \end{pmatrix}, \tag{11}$$

assuming a (p,β)-grid of the size $n = lm$ with, respectively, l and m equally spaced p- and β-values, and we find the least-squares solution of $\tilde{M}w = \tilde{C}$ with the constraint that each element of w be larger than or equal to zero. The extended vector $\tilde{C}$ due to the regularization ensures that there are always more equations than unknowns regardless of the number of original data points in C. Due to the instability of the problem, the direct unconstrained matrix solution would lead to negative values, but in for example the Matlab environment, the positivity constraint is simple to enforce with a standard function. We found that this is more practical than nonlinear optimization with, for example, $w_i = \exp(z_i)$.

We emphasize here that, despite the similar fitting procedures, CDFs are quite different from lightcurves as data. First of all, noise does not show as signal deviations because CDFs are monotone curves. The error in the observed CDF curve is essentially due to convolution (the distribution function is multiplied by the error probability function under the CDF integral), causing the smoothing of the curve (very noisy data would produce a featureless CDF resembling a step function). In our analysis, we do not attempt to deconvolve the original CDF, since the convolution (i.e. error) function is not known: in addition to the random and systematic brightness errors, it depends on the number and temporal distribution of the data points. Also, the addition of more measurements does not fill the gaps between points as in lightcurves: it alters the whole shape of the CDF. A visually good density of points in a CDF does not imply that its shape is near the correct theoretical one from infinitely many points. Computationally, one can use very high densities for interpolating between the actual CDF values to construct the values in the data vector. Then one does not have to use the full high number of observations, which may be helpful for software dealing with the positivity constraint of the solution based on Eq. (11).

Obviously p-values lower than about 0.4 start to become unrealistic, so one could also use an additional regularization function and a lower limit on p-values. We have, however, used the whole scale since the ostensibly unrealistic p-values are usually not heavily occupied and may carry information. For example, especially for smaller asteroids, the small p-values may also indicate irregular shapes (and the shadowing effects of nonzero solar phase angles) or an otherwise increased "noise level" due to systematic and modelling errors. In any case, the CDF method is meant to give a quick overview of a population instead of a detailed portrait, so trying to extract information via prior constraints is not a key concept here. An abundance of small p-values may also indicate that the data are simply not sufficient or otherwise suitable for a reliable result, so such a warning should not be suppressed by regularization.

4. Simulations

To assess the performance of our method with actual data and to check the effect of the simplified model, we perform several simulations. In Sect. 4.1, we explain the setup for our simulations and experiment with the synthetic data, giving graphical presentations of both actual and computed (p,β) solutions. In Sect. 4.2, we discuss ways to apply a post-solution "deconvolution" in order to correct the systematic errors in the solution; the corrected solutions are graphically presented as well.

4.1. Synthetic data

Since the modelling errors and sampling effects in CDF construction dominate over the noise of original brightness data and the CDF errors do not show as signal noise, no standard error estimates are available. Therefore, the only way to test the reliability of our method is via simulations. In our setup, we utilize synthetic data for brightness measurements and attempt to reconstruct the (p,β) distributions. In the simulations with the synthetic data, we use the same geometries (i.e. the direction vectors $\boldsymbol{e}_{\rm sun}$ and $\boldsymbol{e}_{\rm earth}$ as seen from the asteroid-fixed frame) and measurement time information that are used in the real asteroid databases, such as the Lowell Observatory database (Bowell et al. 2014), the Asteroid Lightcurve Data Exchange Format (ALCDEF; Warner et al. 2011), the Uppsala Asteroid Photometric Catalogue (UAPC; Lagerkvist et al. 1987; Piironen et al. 2001) and the Wide-Field Infrared Survey Explorer (WISE; Mainzer et al. 2011).

In our forward model, we take asteroid models from DAMIT and apply basic transformations such as stretching on them to obtain a desired shape distribution with a large number of objects. For DAMIT shapes, the concept of elongation is no longer as well defined as for ellipsoids, but we estimate $p = b/a$ simply by choosing a to be the longest diameter in the equatorial xy-plane, and b the width in the corresponding orthogonal direction. In this way, we generate a (p,β) distribution with one peak by choosing suitable values of p and β for the objects.

For computing the brightnesses of the synthetic asteroids, we use a combination of the Lommel–Seeliger and Lambert scattering laws as in Kaasalainen & Lamberg (2006). We also add random perturbations to L to simulate noise. When the brightness function has been computed, η can be obtained using the discrete approximation of $\Delta(L^2)/\langle L^2\rangle$ from the available synthetic data points close enough in time to depict one rotation in a fixed geometry. Then, we get A from Eq. (8), and thus, the CDF of A: that is, $C(A)$. Other scattering models such as Hapke's could be used as well, but this represents only small brightness changes to separate objects and is thus not relevant to the collective results here. In fact, asteroid models in DAMIT are mostly constructed using the combined Lommel–Seeliger and Lambert law, so this choice reproduces the typical observed asteroid brightnesses best in this simulation.

In the inverse problem, we attempt to reconstruct the original distribution. First of all, the competence of our method depends on how close our obtained distribution is to the original one. We can check numerically how well the function series $\sum_{ij} w_{ij} F_{ij}(A)$ of Eq. (5) converges to the CDF of A, $C(A)$, by computing the relative error:

$$\frac{\left\| C(A) - \sum_{ij} w_{ij} F_{ij}(A) \right\|}{\|C(A)\|} \overset{\text{Eq. (9)}}{=} \frac{\|C - Mw\|}{\|C\|}. \tag{12}$$

Figure 3 depicts a typical fit of the analytical basis functions F_{ij} to the data created with complex shapes and sampled more sparsely than implicitly assumed by the CDF integrals. We can see that the model usually fits CDF data perfectly, so the analytical basis functions provide a very good set despite the crudeness of the model approximations. The main question is thus the accuracy of the result rather than the explainability of the model.

In our simulations with synthetic data, we generate populations with a single (p,β) peak in their joint distribution, and we attempt to reconstruct this peak. Each asteroid can have multiple brightness measurements, and we require at least five observation points for a valid estimate of the variation observable η. From the results of the simulations, we have found this to be

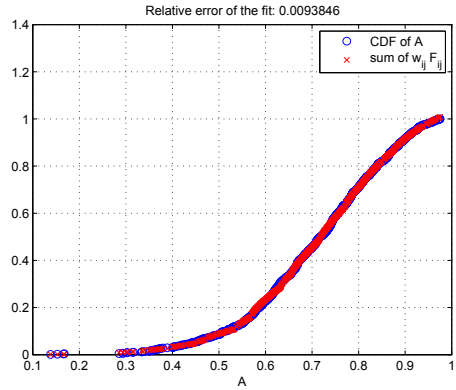

Fig. 3. Function series $\sum_{ij} w_{ij} F_{ij}$ from Eq. (5) plotted in the same figure with the CDF $C(A)$ from data. The minimal error in the fitting should be noted.

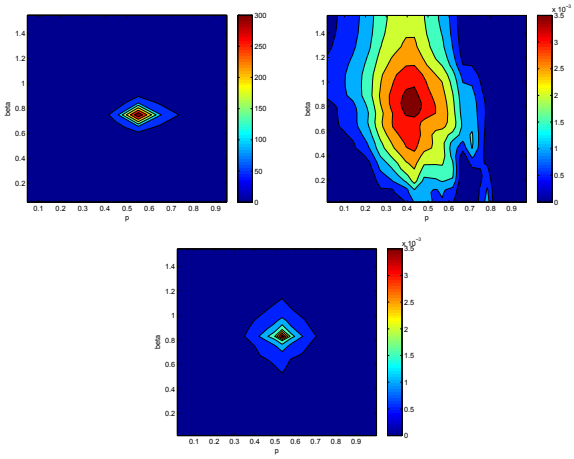

Fig. 4. Actual joint distribution of (p,β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). The colours depict the occupation number of each (p,β)-cell (on arbitrary scales). The absolute value of the ecliptic latitude of the spin axis decreases from bottom (perpendicular to the ecliptic plane) to top (in the ecliptic plane), and the shape elongation decreases from left (thin cigar) to right (sphere). On the *bottom* is the solution of the inverse problem after applying deconvolution for correction.

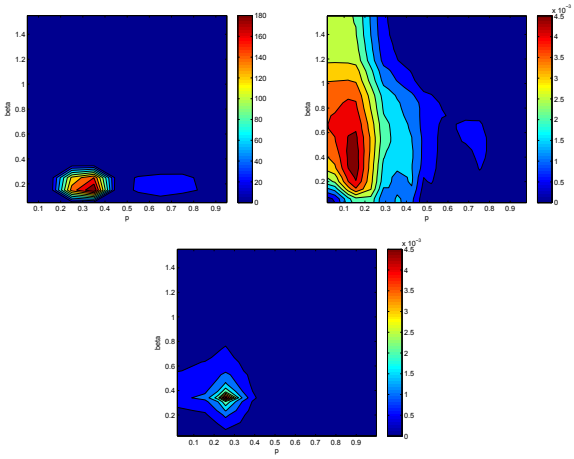

Fig. 5. Actual joint distribution of (p,β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). Here we have tested how accurately the solution is obtained if the peak of both p and β distributions is low. On the *bottom* is the solution with deconvolution added.

the typical minimum number of data points for sampling one rotation of the target. This is also simple to estimate analytically by considering the possible permutations of random samples of a boxcar-shaped sinusoidal signal: then the average error of η drops fast from the 50% of two sample points to close to 10% with five or six points. Such error levels already fit well in the total error budget. Numerical examples of lightcurves give similar results.

In order to obtain enough observations of η and thus accurate distributions, we use populations of 1000 asteroids. This is a realistic population size, as real major databases contain $\propto 10^3$–10^4 objects. Smaller populations start to suffer from too sparse sampling of geometries. Also, the systematic errors caused by the inaccurate assumptions of the model and the sparsity of rotational phases in η estimation are best counteracted by the averaging effect of a large number of samples. For the bins, $p_i \in [0,1]$ and $\beta_j \in [0,\pi/2]$, where we have selected $i = 1, \ldots, 20$ and $j = 1,\ldots,29$, with a random point near the centre of each equally spaced bin to represent its β or p value. This way, every bin is about 0.05×0.05 units in size, and the computation of the inverse problem is fast enough.

We have plotted some all-round cases of the (p,β) distribution of the forward model and the solution distribution of the inverse problem (from data mimicking WISE) in Figs. 4–8. In the forward model, the peak of the (p,β) distribution has been placed in the middle, bottom left, top left, bottom right, and top right positions in the (p,β) plane, respectively. In Fig. 4, we notice that the approximate location of the (p,β)-peak is correct, but the solution spreads when moving away from the peak, particularly when moving towards $\pi/2$ (spin direction in the ecliptic plane). In addition, the peak is too much to the left in p-axis (towards more elongated bodies). The same phenomenon can be observed in the other figures as well. In Fig. 6, when the peak was located in the top left corner of the (p,β) plane (elongated bodies in the ecliptic plane), the contour looked visually messy every time. Hence, we included an additional plot of the solution of the inverse problem in the $(p,\beta,\mathrm{DF}(p,\beta))$ coordinates. From the three-dimensional perspective, we notice that the peak of the shape elongation is once again too far in the left in p-axis, and the solution spreads when moving away from the peak. Indeed, these errors are systematic, and they occur in a solution every time. The p-shift is inevitable: even in the absence of noise, near-spherical targets with $p = 1$ do not portray a completely flat lightcurve because of local shape irregularities. In addition, Figs. 5, 7, and 8 show a trend of the peak of the β solution to move slightly towards the middle (away from the $\beta = 0$ and $\beta = \pi/2$ ends). The error is common but does not occur every time. In general, the errors are encountered because of both modelling errors and noisy measurements. They are rather regular and predictable, and therefore, it is possible to formulate a post-solution correction in order to revise the solution distribution.

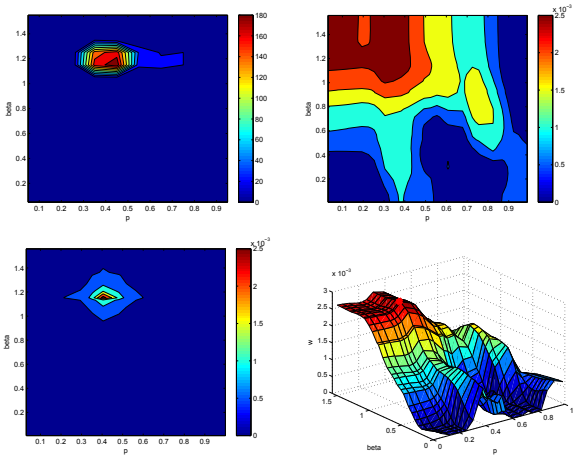

Fig. 6. Actual joint distribution of (p,β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). To make the solution plot more easily readable, we have plotted it from another perspective on the *bottom right*, with the z-axis depicting the weights w of each (p,β) bin. Here we have tested how accurately the solution is obtained if the peak of the distributions is low for p and high for β. On the *bottom left* is the solution with deconvolution.

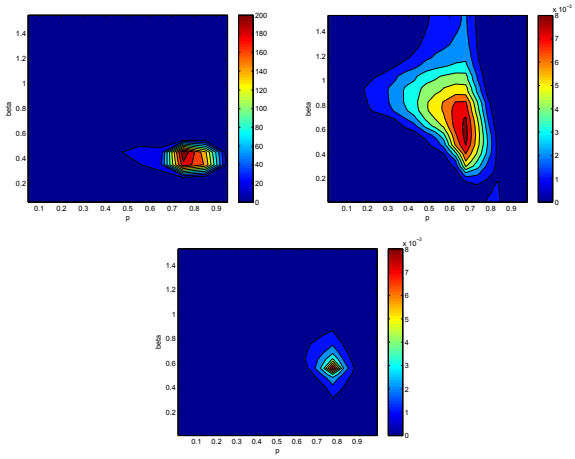

Fig. 7. Actual joint distribution of (p,β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). Here we have tested how accurately the solution is obtained if the peak of the distributions is high for p and low for β. On the *bottom* is the solution with deconvolution.

4.2. Correction in (p, β)-plane

The "deconvolution" of the noisy solution "image" in the (p,β)-plane is a visual aid based on experiments performed on the synthetic data. With the help of simulations, we were able to acquire a good understanding of how much our computational solution typically differs from the actual distribution when one assumes that there is a dominant peak in the latter. This way, we could deduce the typical point-spread function of the solution in the plane. We introduce damping on bins further away from the peak of the centre of the solution. Then we move the values of p a constant (fixed) step to the right: $p_i \rightarrow p_i + \Delta\mathcal{P}$. According to our simulations with the WISE database, the solution is typically

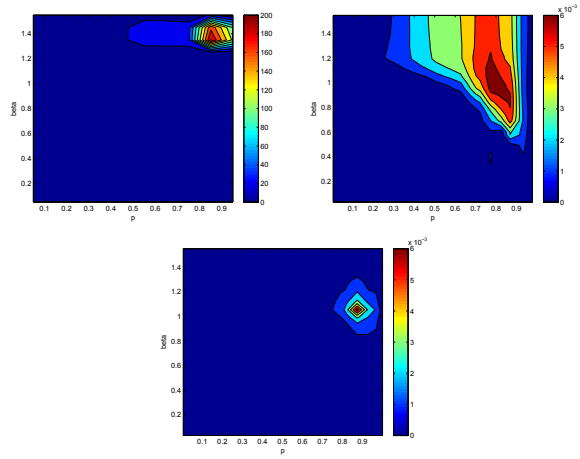

Fig. 8. Actual joint distribution of (p,β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). Here we have tested how accurately the solution is obtained if the peak of both p and β distributions is high. On the *bottom* is the solution with deconvolution.

shifted about 0.1 p-units to the left due to noise. Therefore, we choose $\Delta\mathcal{P} = 0.1$ when we use WISE data. The systematic error in β direction is irregular, and there is no way to know whether the obtained β is too small or too large. For β, we observed that the obtained distribution is usually accurate if the actual β peak is somewhere near $\pi/4$ (usually when $\beta \in [0.5, 1]$), but if the actual peak is near the extreme end values 0 or $\pi/2$, then the solution tends to shift the peak away from the extremes, towards the middle values.

We show the deconvoluted solution in the bottom figures of Figs. 4–8 (bottom left picture in Fig. 6). The deconvolution has been used to correct the solution presented in the top right picture of the same figure. The corrected solution is close to the distribution shape of the forward model in the top left picture. We will apply deconvolution solely on the joint (p,β) distribution. In order to reduce errors and loss of information, the marginal p and β distributions are presented without corrections. In their cases, the main point is that the peak of the p distribution is usually slightly more to the right in the p-axis than in the obtained solution.

5. Results from astronomical databases

In this section, we plot distributions of different asteroid families and introduce a method for comparing such distributions. The setup we use is very similar to the one used for synthetic data. We receive our data (geometries, brightness values, measurement times) from the WISE database. We downloaded the data from the Infrared Science Archive: infrared Processing and Analysis Center (IRSA/IPAC archive[4]) and used the same selection criteria as Ali-Lagoa et al. (2014). The combined ALCDEF & UAPC lightcurve database (hereafter called simply ALCDEF) is also useful for various analyses, but its denser lightcurves (yielding improved η estimates) do not really compensate for the larger number of objects in the WISE data that is crucial to the robustness of the statistical CDF approach, as discussed earlier. Moreover, the lightcurves sample well only asteroids with short

[4] http://irsa.ipac.caltech.edu/Missions/wise.html

rotation periods, because observations from different nights usually cannot be combined together because of poor or completely missing calibration. The large number of brightness variation samples, rather than the accuracy of the observable, is the main reason why the method can tolerate the crude underlying shape model. Even though the WISE data are in mid-infrared wavelengths (we used measurements at 12 and 22 μm), the η derived for them is essentially the same as from the projected area since the infrared regime mainly causes a lag in the lightcurves (compared to visual data) that does not affect the brightness variation (Ďurech et al. 2016a).

Another possible rich source of asteroid photometry is the Lowell Observatory photometric database Bowell et al. (2014), which was used by Oszkiewicz et al. (2011) and Cibulková et al. (2016). However, the large errors of photometric points of ~0.1–0.2 mag would bring another source of systematic error into our model.

We consider one η-estimate to consist of measurements done within a three-day time window to keep the observing geometry sufficiently constant. We only accept estimates based on at least five measured values. In principle, our model requires the phase angle between the Sun and Earth to be close to zero degrees. However, adding this restriction would greatly reduce the number of brightness measurements we could use, and would eventually lead to a considerably lower number of η values. According to simulations performed on synthetic data, the error caused by a non-zero phase angle is so small compared to other error sources that its effect is negligible. Therefore, we set a very liberal requirement that

$$\arccos(\boldsymbol{e}_{\rm sun} \cdot \boldsymbol{e}_{\rm earth}) \leq 30^\circ. \quad (13)$$

5.1. Discussion about bias

Before we move on to plotting asteroid families, we discuss some possible sources for biases. The number of possible η-estimates varies between asteroids; if an asteroid yields n estimates of η, we can formally give each of these estimates the weight $1/n$. However, if there are many estimates associated with some asteroids, there could be a bias, as the solution could be favouring such targets. We checked if the solution was affected if we only took one estimate for each target. There was no noticeable change from the situation when all estimates were considered, so we can conclude that there is no significant bias from the weights of individual asteroids when using large databases for a large number of objects. This underlines the safety in large numbers, so the method can be used even if we do not know which observation is from which asteroid and use all η estimates "blindly".

As we know from asteroid lightcurves, large asteroids are generally more spherical than small asteroids. This was also shown by Cibulková et al. (2016), for example. We checked this result by comparing WISE subpopulations of different sizes. We divided the WISE asteroids into four subpopulations: ones with diameter $D < 10$ km (≈65 000 bodies), 10 km $\leq D < 25$ km (≈6000 bodies), 25 km $\leq D < 50$ km (≈1000 bodies), and $D \geq 50$ km (≈1000 bodies). While the population sizes are noticeably different (the number of small asteroids clearly surpasses the number of large ones), the subpopulations were selected so that each of them would have a sufficiently large sample size in order to acquire reliable results. The obtained distributions confirmed that large asteroids indeed tend to be more spherical (see Fig. 9 for comparison). Our result is considerably different from the one obtained by McNeill et al. (2016), where the peak of the p distribution for $D < 8$ km was located at a near-spherical value of $b/a = 0.85$. Of course, the distribution tail of small p-values of our result for $D < 10$ km is more due to systematic and model errors (especially irregular shapes) than to actually very elongated bodies.

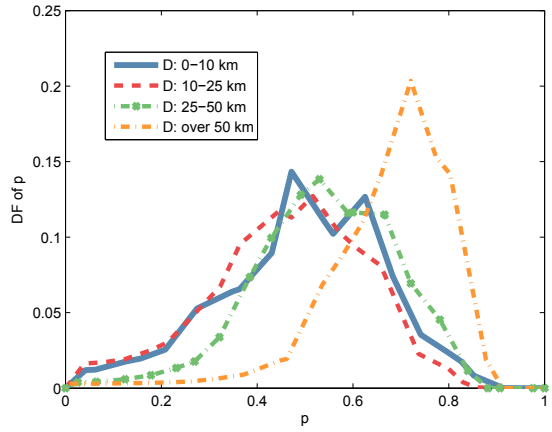

Fig. 9. Comparison of the marginal DFs of the shape elongation p for asteroids of different sizes from WISE data. The vertical axis depicts the occupation number of each p-slot (on an arbitrary scale). The shape elongation decreases from left (thin cigar) to right (sphere).

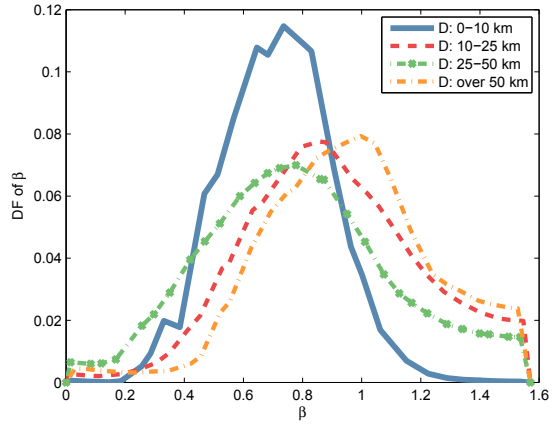

Fig. 10. Comparison of the number densities of the spin β for asteroids of different sizes from WISE data. The absolute value of the ecliptic latitude of the spin axis decreases from left (perpendicular to the ecliptic plane) to right (in the ecliptic plane).

Similarly, large asteroids tend to have their spin axes closer to the ecliptic plane than the small ones (see Fig. 10), which qualitatively agrees with the results of Hanuš et al. (2011) and Ďurech et al. (2016b), who studied the distribution of spins of asteroids on a sample of several hundred individual models. Contrary to the results based on individual models where the spins are clustered towards poles of ecliptic ($\beta = 0$ in our notation), our analysis shows that small asteroids have a remarkably sharp peak extending some 15° on both sides of the ecliptic spin latitude of 50°, but the lack of values $\beta \approx 0$ might be caused by some systematic effects of our simple model (but we note the effect of $\sin\beta$ in the plot as discussed in Sect. 2.1).

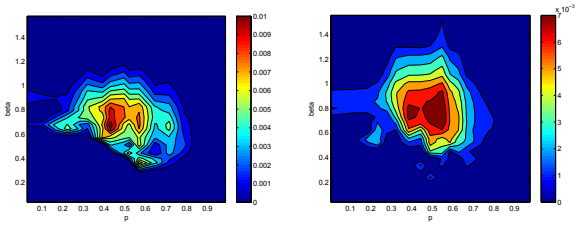

Fig. 11. Contour solution of the joint (p,β) distribution based on WISE (*left*) and WISE+ALCDEF data (*right*). Deconvolution is not used to preserve the double-peak information.

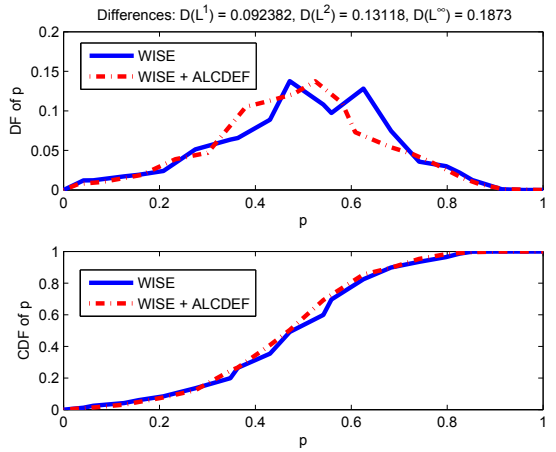

Fig. 12. Comparison of the marginal DFs (*top*) of the shape elongation p for the entire WISE population and merged ηs from WISE and ALCDEF databases, and of their marginal CDFs (*bottom*).

When comparing asteroid families, the size of the targets may thus be a factor. The shape and/or spin difference between two families may be partly driven by the difference in their size distributions. On the other hand, the best statistical material is acquired without adding the size as another dimension in the distribution function since there usually are just not enough targets to split a family into size bins. Thus we report the family p and β distributions here as such, using all family members without considering the size distribution a bias factor, even though a closer analysis between families may require taking at least the division into small and intermediate sizes into account. The proportion of large asteroids is small, so their contribution to the population distributions is usually small as well. As an example, we investigated the Eos family of well over 3000 η estimates. We performed some comparisons of p and β distributions for the whole family and its subset of smaller bodies with a diameter less than 20 km. For the p distribution, the inclusion of large asteroids mainly affected the width of the peak and there was no noticeable difference in the distributions. The differences were even smaller for the β distribution. These results suggest that the biases caused by large objects are insignificant and we will include full populations in our examples. For this and other comparison purposes, we propose a measure of difference tailored to our case (instead of the Kolmogorov–Smirnov test that typically produces indecisive statistics).

Let S_1 and S_2 be two sets of population samples. For S_i, let $F_p(S_i)$ and $F_\beta(S_i)$ be the CDFs for the marginal distributions of DF solutions for p and β, respectively (normalized to the interval $[0,1]$). We define the statistical difference measure between S_1 and S_2 as

$$\begin{cases} D_p(S_1,S_2) = \alpha_k \left\| F_p(S_1) - F_p(S_2) \right\|_k \\ D_\beta(S_1,S_2) = \alpha_k \left\| F_\beta(S_1) - F_\beta(S_2) \right\|_k , \end{cases} \tag{14}$$

where usually $k = 1$, $k = 2$ or $k = \infty$, and α_k is a norm-based scaling factor to fix the statistical difference to the same magnitude for all norms; typically, $\alpha_1 = 1/4$, $\alpha_2 = 1$ and $\alpha_\infty = 2$. The case $k = \infty$ is used in the Kolmogorov–Smirnov test. In addition to the L^∞ norm, we will also compute the L^1 and L^2 norms, as this way we will have a better understanding of the type of statistical difference, for example, do the distributions differ in terms of the maximum or mean difference. Generally, our simulations suggest that in this context two distributions can be considered statistically different if $D \gtrsim 0.2$, although one number does not tell the whole story, and it is more instructive to perform a visual inspection on the marginal DF and CDF plots.

In Eq. (14), we chose to compare the CDFs rather than the DFs since the latter is the derivative of the former and CDFs are monotone functions, so computing the norm of their differences is more stable when one aims at one number depicting the difference. The CDF difference tells whether the distributions are really different in the first place, and the DFs give additional details of the potential differences.

Finally, we ran tests to see if there is a bias associated with the selected database. WISE is one of the largest databases, with our method being able to cover about 85 000 asteroids. We can get about 86 000 values of η from WISE, which means we can get approximately one value for our observable from each asteroid. Indeed, WISE is one of the biggest asteroid databases available. The average number of brightness measurements available for one η is $\langle n_L \rangle \approx 9$. The ALCDEF lightcurve database contains about 14 000 asteroids suitable for our method, and yields about 39 000 values of η, resulting in less than three values of η from each asteroid. For ALCDEF, $\langle n_L \rangle \approx 18$. Despite its smaller sample of asteroids, the ALCDEF's η-per-asteroid ratio is better than WISE's. The drawback of the ALCDEF is the weak or nonexistent calibration as well as selection effects.

We perform a consistency check by comparing two distributions. First, we compute the solution of the inverse problem using all ηs obtained from the WISE population. Then we do the same, but using all the ηs from both WISE and ALCDEF. The (p,β)-plane plots are shown in Fig. 11.

The p distributions of WISE and WISE+ALCDEF are plotted in Fig. 12, while the β distributions of WISE and WISE+ALCDEF are plotted in Fig. 13. We have not used deconvolution procedures since, being designed for one-peak distributions, they would smooth out the two-peaked result. For p distributions, a visual inspection shows some differences between the DFs, such as the unrealistic boost of p-values around 0.4 by the ALCDEF addition. The CDFs are similar enough to suggest that the added ALCDEF data do not greatly distort the p distribution. For β, on the other hand, the addition of the ALCDEF data shifts the distribution to the right, and the β distribution obtained from the hybrid data is obviously different from the one with WISE data only. Therefore, we conclude that there may be a database-related bias included, especially with the β solutions, and it is advised to use caution in the selection of a database. Indeed, the ALCDEF data are distorted by a number of selection effects due to the visibility and popularity of the targets. WISE targets are more evenly and comprehensively spread and

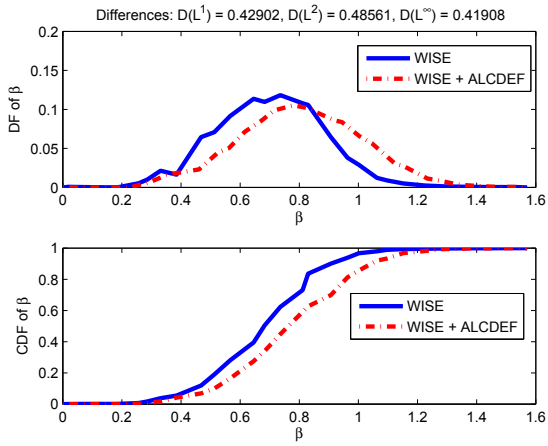

Fig. 13. Comparison of the marginal DFs (*top*) of the spin β for the entire WISE population and merged ηs from WISE and ALCDEF databases, and of their marginal CDFs (*bottom*).

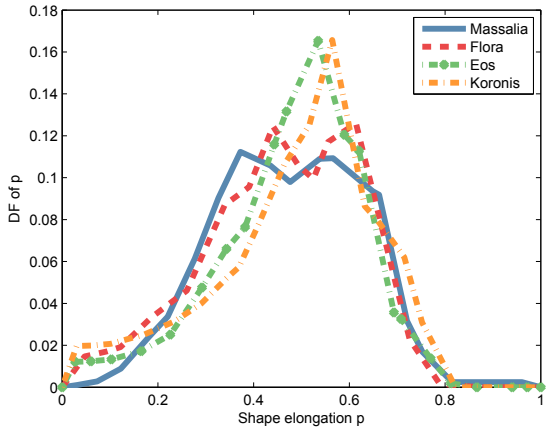

Fig. 14. Obtained p distributions for the asteroid families Massalia, Flora, Eos, and Koronis.

observed from a satellite, so the biases are smaller. Due to this and the sufficiently large number of database targets, we consider WISE data more reliable for distribution analysis and use them in the studies below.

5.2. Examples of distributions and their comparison

Below we list some results from family distribution reconstruction and comparison by their inferred marginal distributions of p and β. We chose families that were interesting from a statistical point of view, mainly to demonstrate the features, differences, and similarities the method can discover. Obviously, in addition to the interpretation and analysis of the results, there are many other families as well as populations other than families to consider in further work from astronomical points of view. We note that the computed distributions for asteroid families vary slightly in the figures, just to illustrate that randomized inverse grids lead to slightly different details in distribution solutions.

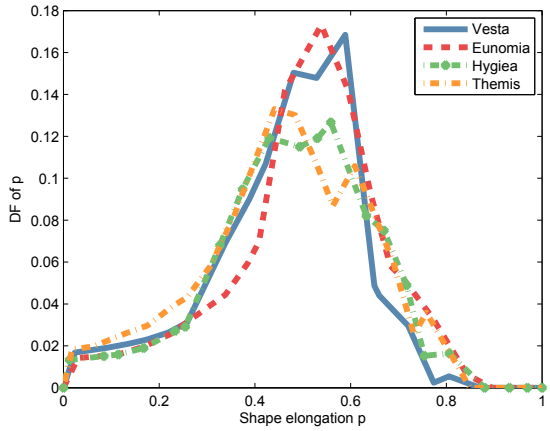

Fig. 15. Obtained p distributions for the asteroid families Vesta, Eunomia, Hygiea, and Themis.

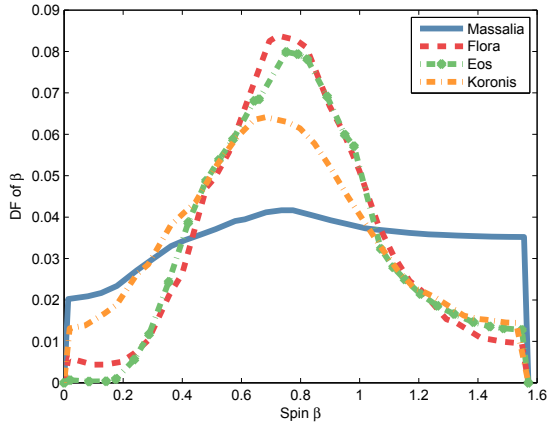

Fig. 16. Obtained β distributions for the asteroid families Massalia, Flora, Eos, and Koronis.

The shape elongation and spin distributions for eight different asteroid families are shown in Figs. 14–17. The obtained shape elongation distributions are different from those obtained by Szabo & Kiss (2008), who assumed a uniform distribution of spin axes and utilized the less reliable two-point brightness scatter observable. The number of available asteroid samples per family in the WISE database varies. Flora and Eos have about 3000 WISE samples, while Vesta, Eunomia, Hygiea, and Themis have 1000–2000. Koronis has 642 WISE samples, while Massalia is limited to only 154 samples. According to our simulations, solving the inverse problem several times for Koronis leads to fairly good regular solutions, so the sample size can be trusted to be large enough. Massalia, on the other hand, has stability problems with the small sample size, so its results cannot be considered to be as reliable as the others, but we include it for completeness. Typically, a sample of at least 500 objects is required in order to obtain stable solutions that can successfully recover from the model errors and noise.

We observed that the Alauda family (some 800 samples) contains a somewhat higher ratio of near-spherical bodies than other families. This is likely to be an intrinsic quality of the family, as the same result holds when the large asteroids ($D > 20$ km) have

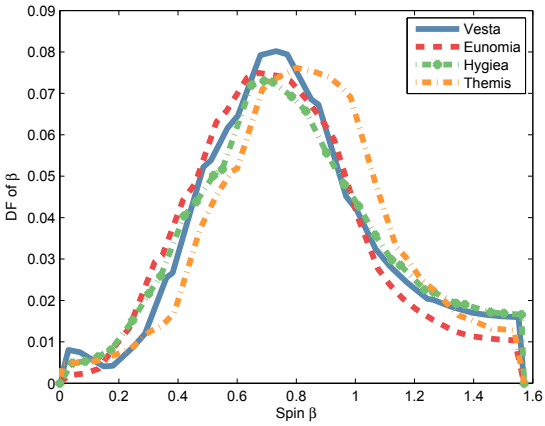

Fig. 17. Obtained β distributions for the asteroid families Vesta, Eunomia, Hygiea, and Themis.

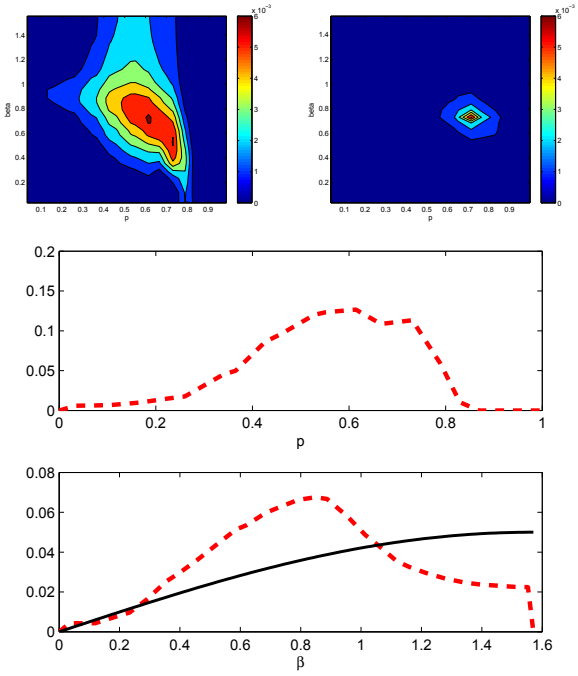

Fig. 18. Contour solution of the joint (p,β) distribution of the Alauda family (*top left*), the deconvoluted smoothing of the same solution (*top right*), and the normalized marginal distributions for p (*middle*) and β (*bottom*). The black solid curve ($\sin\beta$) depicts the curve shape of a constant level of spin distribution on the sphere. Deconvolution is not used for the marginal distributions in order to avoid the loss of information in the smoothing.

been filtered out. The contour solution of the joint (p,β) distribution, both with and without deconvolution, as well as the marginal p and β distributions (without deconvolution) for the Alauda family are plotted in Fig. 18. We plotted an additional $\sin\beta$ curve in the β plot to illustrate what the spin DF would look like if it was uniformly distributed on the sphere.

To show examples of difference classes between families, we consider cases of small, borderline, and large values of the

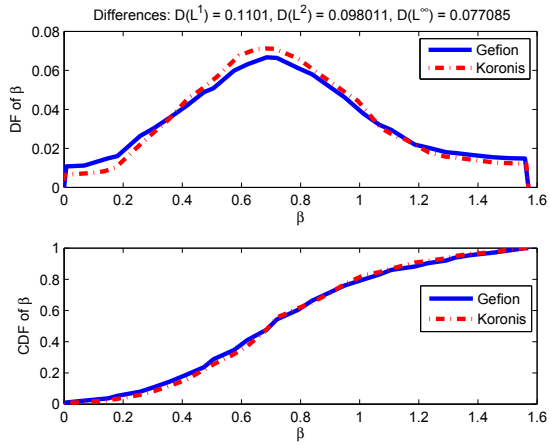

Fig. 19. Comparison of the marginal DFs (*top*) of the spin β for Gefion and Koronis families, and of their marginal CDFs (*bottom*).

difference measures. The very similar β distributions of the Gefion and Koronis families are shown in Fig. 19, while the β distributions of the Phocaea (some 1000 samples) and Alauda families are plotted in Fig. 20. Phocaea has slightly more asteroids with spins closer to perpendicular to the ecliptic plane, while both families have heavy tails close to the plane. Generally, our solutions appear to give less weight to the values of β close to 0 or $\pi/2$. In the case of $\beta = 0$, this is partly due to the factor of $\sin\beta$ in the occupation numbers. For $\beta = \pi/2$, this can be due to, for example, size distribution, noise, and orbit positions away from the ecliptic plane. In our simulations, we did not find any particular mechanism or tendency for the scarcity of solutions close to the ecliptic plane. Finally, in Fig. 21, we give an example of the clearly different shape distributions of the Themis and Alauda families.

To confirm the reliability of distributions of β and p for individual families, it would be ideal to compare our CDFs with those constructed from individual models derived by lightcurve inversion. Unfortunately, this is not possible at this stage, because the number of known models for a typical family is a few tens at most (Hanuš et al. 2013). Another possibility would be to compare CDFs reconstructed from different and independent data sets. For example, the difference between Themis and Alauda families is also significant when we do the same analysis with the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) data (Cibulková et al., in prep.).

6. Discussion and conclusions

The statistical CDF approach is a fast way of testing hypotheses about shape and spin distributions of asteroid populations without constructing models of separate objects. It is applicable for discovering the existence of peaks in distributions of parameters and for comparing the distributions of different populations.

There are numerous possibilities of analysing and comparing asteroid families, and a comprehensive analysis and interpretation of results is not the aim of this paper. Our main goal was to discuss the usefulness and various aspects of the CDF approach, and build mathematical tools for its efficient use. As we have seen, the data and inversion procedures of the problem are, in fact, very simple and fast to generate and apply as such. The main burden lies in the judicious interpretation of the

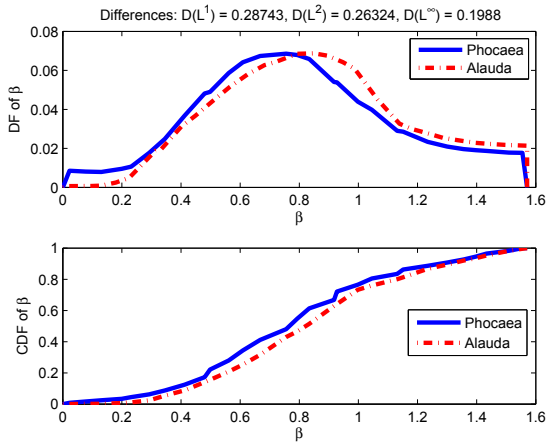

Fig. 20. Comparison of the marginal DFs (*top*) of the spin β for Phocaea and Alauda families, and of their marginal CDFs (*bottom*).

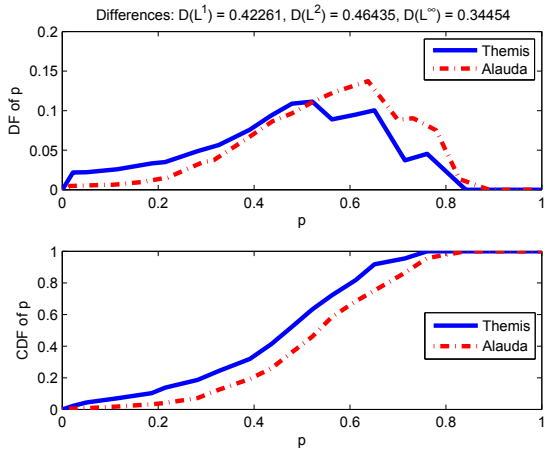

Fig. 21. Comparison of the marginal DFs (*top*) of the shape elongation p for Themis and Alauda families, and of their marginal CDFs (*bottom*).

results. A scrutiny of the usable databases may well be necessary, as some data sources may cause skewed results due to biases and/or noise.

We introduced a robust observable that can provide information on both the shape elongation and spin properties of asteroid populations, and performed an analysis on the theoretical background, providing also some examples of the obtained distributions, inspected from the statistical point of view. In our analysis, we proved that unique solutions can be obtained for both shape elongation and spin distributions, and we performed numerical simulations in order to verify that they coincide with the analytical results. It is interesting to note that, while the abundance of orbits close to the ecliptic plane means that many individual asteroid models necessarily have an ambiguity of 180 degrees in the spin longitude (Kaasalainen & Lamberg 2006), the same ecliptic orbital configuration makes possible the population-level information on spin latitudes.

Due to the model noise and the assumptions made, we cannot expect to obtain detailed, high-resolution solutions of the distributions, but we get the overall picture when the observational noise is sufficiently low (at most 0.05 mag or so). If the data noise is large, the whole point of using the brightness variations as the observable is challenged, and the prior information needed for regularization would dominate the solution. Typically, high noise means that the real p-information on near-spherical bodies disappears, and the β-information is similarly severely diluted. For low observational noise, we can additionally use a deconvolution filter to correct the systematic errors caused by modelling errors and noise. The deconvolution is a visual tool which attempts to illustrate what the distribution of the parameters actually looks like, based on prior information obtained from simulations performed for synthetic data.

Trans-Neptunian objects (TNOs) are especially interesting from the statistical point of view due to their slow orbital motion. Since their observing geometries change very slowly, proper individual models of TNOs cannot be made with the current ground-based instruments in less than tens of years. Also, for TNOs the solar phase angle is essentially zero, leading to ambiguous shape solutions from photometry. The statistical approach, however, applies to TNOs just as well as to other populations so long as there are sufficient targets in the observed set. This is, in fact, the only way to model TNO populations with data from large-scale surveys. A bonus with TNOs is that, because of the essentially fixed geometry and near-zero solar phase angle, all calibrated survey data points are usable for η-estimation even if they are separated by long time intervals. With main-belt asteroids, sparse data points from one apparition may be usable together for η estimates if one uses a solar phase correction as in Kaasalainen (2004), Ďurech et al. (2009), or Cibulková et al. (2016). The additional systematic error from this is not necessarily very large considering the total error budget.

We emphasize that the statistical use of brightness variation, while a promising approach, should always be treated with caution. Above all, any proposed type of observable and implementation should be checked with realistic simulations where the synthetic data are created with a model different from the one used in inversion. The information potential of each dataset should be assessed by using its actual observing geometries in the simulations. These simulations yield insight into the uniqueness and stability properties and accuracy expectations. Based on analytical considerations and simulations mimicking real databases, we advocate the use of the η observable and the corresponding analytical basis functions in the inverse problem.

We plan to offer a software application as a statistical (and simulation) tool that can be used for experimenting with different populations that are defined by the user. It may also be useful to construct solution procedures tailor-made for input populations. For example, one can create basis functions numerically by making synthetic CDFs for each (p,β)-bin with DAMIT-based shapes placed in the orbits of the populations.

Acknowledgements. We would like to thank Matti Viikinkoski for valuable comments and discussions as well as assistance with software. This research was supported by the Academy of Finland (Centre of Excellence in Inverse Problems), and H.N. was supported by the grant of Jenny and Antti Wihuri Foundation. J.Ď. and H.C. were supported by the grant 15-04816S of the Czech Science Foundation. VAL has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements Nos. 640351 and 687378. This publication also makes use of data products from NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the Planetary Science Division of the National Aeronautics and Space Administration. In addition, this research made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory/California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

Ali-Lagoa, V., Lionni, L., Delbo, M., et al. 2014, A&A, 561, A45

Bowell, E., Oszkiewicz, D. A., Wasserman, L. H., et al. 2014, Meteoritics and Planetary Science, 49, 95

Cibulková, H., Ďurech, J., Vokrouhlicky, D., Kaasalainen, M., & Oszkiewicz, D. 2016, A&A, 596, A57

Connelly, R., & Ostro, S. 1984, Geometriae Dedicata, 17, 87

Ďurech, J., Kaasalainen, M., Warner, B., et al. 2009, A&A, 493, 291

Ďurech, J., Carry, B., Delbo, M., Kaasalainen, M., & Viikinkoski, M. 2015, in Asteroid Models From Multiple Data Sources, in Asteroids IV, eds. P. Michel, et al. (Tucson: University of Arizona Press), 183

Ďurech, J., Hanuš, J., Ali-Lagoa, V., Delbo, M., & Oszkiewicz, D. 2016a, WISE data and sparse photometry used for shape reconstruction of asteroids, in Asteroids: New Observations, New Models, Proceedings of the International Astronomical Union, IAU Symp., 318, 170

Ďurech, J., Hanuš, J., Oszkiewicz, D., & Vančo, R. 2016b, A&A, 587, A48

Hanuš, J., Ďurech, J., Brož, M., et al. 2011, A&A, 530, A134

Hanuš, J., Brož, M., Ďurech, J., et al. 2013, A&A, 559, A134

Kaasalainen, M. 2004, A&A, 422, L39

Kaasalainen, M., & Lamberg, L. 2006, Inverse Problems. 22, 749

Lagerkvist, C., Barucci, M. A, Capria, M. T., et al. 1987, Asteroid photometric catalogue, Istituto di Astrofisica Spaziale, Consiglio Nazionale delle Ricerche

Mainzer, A., Bauer, J., Grav, T. et al. 2011, ApJ, 731, 53

McNeill, A., Fitzsimmons, A., Jedicke, R., et al. 2016, MNRAS, 459, 2964

Oszkiewicz, D. A., Muinonen, K., Bowell, E., et al. 2011, AAPP, 89, C1V89S1P072

Piironen, J., Lagerkvist, C., Torppa, J., Kaasalainen, M., & Warner, B. 2001, in BAAS, 33, 1562

Szabó, G., & Kiss, L. 2008, Icarus, 196, 135

Warner, B. D., Stephens, R. D., & Harris, A. W. 2011, Minor Planet Bulletin, 38, 172

Appendix A: CDF integrals and analytical basis functions

We recall the expression of the amplitude A, given by Eq. (2):

$$A = \sqrt{1 + (p^2 - 1)\sin^2\theta}.$$

From this expression, the curves of constant A in the (p, θ)-plane are given by

$$\cos^2\theta_A(p) = \frac{A^2 - p^2}{1 - p^2} := g_A(p). \tag{A.1}$$

The solutions for θ_A are convex "ripples" starting from the point $(p = 0, \theta = \pi/2)$ (upper left corner) for $A = 0$ and continuing to the lines $\theta = 0$ and $p = 1$ for $A = 1$ (lower right corner). Denoting the model DF of elongation by $f(p)$, we write the unnormalized CDF $C(A)$ as

$$C(A) = \int_0^{p_{\max}(A)} f(p) \int_{\theta_A(p)}^{\pi/2} \sin\theta \,\mathrm{d}\theta \,\mathrm{d}p,$$

where the minimal shape elongation needed to produce amplitude A, obtained at $\theta = \pi/2$, is $p_{\max}(A) = A$. With a change of variable $x = \cos\theta$, we get

$$C(A) = \int_0^A f(p) \int_0^{\sqrt{g_A(p)}} \mathrm{d}x \,\mathrm{d}p = \int_0^A f(p)\sqrt{g_A(p)}\,\mathrm{d}p. \tag{A.2}$$

We can also include the effect of spin distribution. Assuming λ to be isotropic and the observation directions to be in the xy-plane of the inertial frame (as they approximately are for the majority of asteroids, when this plane is that of the Earth's orbit), we study the DF $f_\beta(\beta)$ (or the joint DF $f(p, \beta)$ with p). The minimal aspect angle is $\theta_{\min} = \pi/2 - \beta$. Now, substituting $\boldsymbol{e} = (\cos\lambda_e, \sin\lambda_e, 0)$ into $\cos\theta = e_1 \sin\beta\cos\lambda + e_2 \sin\beta\sin\lambda + e_3\cos\beta$, we have

$$\cos\theta = \sin\beta\cos\Lambda,$$

where $\Lambda := \lambda - \lambda_e$ is assumed isotropic (evenly distributed longitudes of spins and observing directions). It is sufficient to explore the region $\Lambda \in [0, \pi/2]$ as other quadrants are just symmetric multiples.

The curves of constant θ

$$\Lambda_\theta(\beta) = \arccos\frac{\cos\theta}{\sin\beta} \tag{A.3}$$

in the (β, Λ)-plane are now expanding "ripples" of increasing θ starting from the point $(\beta = \pi/2, \Lambda = 0)$ for $\theta = 0$. The CDF for θ is, with $x = \cos\beta$ (but retaining the argument β in the DF for convenience),

$$\begin{aligned} C_\theta(\theta) &= \int_{\pi/2-\theta}^{\pi/2} f_\beta(\beta)\sin\beta \int_0^{\Lambda_\theta(\beta)} \mathrm{d}\Lambda \,\mathrm{d}\beta \\ &= \int_{\pi/2-\theta}^{\pi/2} f_\beta(\beta)\sin\beta\Lambda_\theta(\beta)\,\mathrm{d}\beta \\ &\overset{\text{Eq. (A.3)}}{=} \int_0^{\sin\theta} f_\beta(\beta)\arccos\frac{\cos\theta}{\sqrt{1-x^2}}\,\mathrm{d}x. \end{aligned}$$

(Differentiating $dC_\theta(\theta)/d\theta$ yields $\sin\theta$ when $f_\beta = 1$ as expected for isotropic spins.)

Using the complement of C_θ (i.e. $\hat{C}_\theta$ in the decreasing direction from $\theta = \pi/2$ to $\theta = 0$) to write the number of states between $\theta_A(p)$ and $\theta = \pi/2$, our CDF $C(A)$ is, analogously with Eq. (A.2) and using Eq. (A.1),

$$C(A) = \int_0^A \left[\frac{\pi}{2}\int_0^1 f(p,\beta)\,\mathrm{d}x - \int_0^{\sqrt{1-g_A(p)}} f(p,\beta)\arccos\frac{\sqrt{g_A(p)}}{\sqrt{1-x^2}}\,\mathrm{d}x\right]\mathrm{d}p. \tag{A.4}$$

(The use of x is merely a matter of convenience for the integration limits.)

The basis functions, that is, any CDF $C(A)$ caused by all objects having given fixed p_i and β_j, are now easy to write in closed form. They are obtained by replacing p and β in the integrands by the fixed p_i and β_j, setting $f = 1$, and using the integration limits to describe the inequalities between A and p, β to define the piecewise function $C(A)$. This replaces the integral by a sum of such basis functions each multiplied by the corresponding weight of the p_i and β_j bin. The resulting basis functions are given in Sect. 2.1.

The above assumption of most orbits to be close to the ecliptic plane is only approximate, and one can always define populations (especially those of near-Earth asteroids) for which it is not true even approximately. Thus the validity of this assumption should be checked for the targets used. However, as we show in Sect. 4, the assumption works quite well (given the large model error budget in any case) with typical asteroid populations for which there is some concentration of viewing geometries sufficiently near the ecliptic plane.

Appendix B: Ill-posedness caused by two-point variation observables

As earlier, we consider the case when θ is (approximately) the same for the pair. Now we have, for two rotation phases ϕ_0 and ϕ,

$$\frac{1 + (p^2-1)\sin^2\theta\cos^2\phi}{1 + (p^2-1)\sin^2\theta\cos^2\phi_0} = q^2,$$

so, with $0 < q \le 1$, that is, $\phi \le \phi_0$ (due to symmetry, we only need to consider the interval $0 \le \phi \le \pi/2$), we define iso-q contours in the (ϕ, θ) plane (for given p, ϕ_0) by

$$r(q, p, \phi_0, \phi) := \frac{q^2 - 1}{(p^2-1)(\cos^2\phi - q^2\cos^2\phi_0)},$$

so, to have viable solutions for θ_q from $\sin^2\theta_q = r$, we must have $p \le q$, $\phi \le \phi_0$, and

$$\cos^2\phi \ge \frac{q^2-1}{p^2-1} + q^2\cos^2\phi_0 := s(q, p, \phi_0) \ge \cos^2\phi_0,$$

so ϕ exist for given p, q, ϕ_0 only if $s \le 1$; that is,

$$\cos^2\phi_0 \le \frac{p^2 - q^2}{q^2(p^2-1)} := t(q, p).$$

Denoting

$$\tilde{s}(q, p, \phi_0) := \arccos\sqrt{s(q, p, \phi_0)}, \quad \tilde{t}(q, p) := \arccos\sqrt{t(q, p)},$$

our CDF is thus

$$\begin{aligned} C_q(q) &= \int_0^q f(p)\int_{\tilde{t}(q,p)}^{\pi/2}\int_0^{\tilde{s}(q,p,\phi_0)}\int_{\theta(q,p,\phi_0,\phi)}^{\pi/2} \sin\theta'\,\mathrm{d}\theta'\,\mathrm{d}\phi\,\mathrm{d}\phi_0\,\mathrm{d}p \\ &= \int_0^q f(p)\int_{\tilde{t}(q,p)}^{\pi/2}\int_0^{\tilde{s}(q,p,\phi_0)} \sqrt{1 - r(q, p, \phi_0, \phi)}\,\mathrm{d}\phi\,\mathrm{d}\phi_0\,\mathrm{d}p. \end{aligned} \tag{B.1}$$

Again, we can include the β-distribution by expanding the integral in the same way as with Eq. (A.4).

The basis function $G_i(q)$ for a given p_i in the two-point brightness scatter case is, from Eq. (B.1),

$$G_i(q) = \begin{cases} 0, & q \le p_i \\ \int_{f(q,p_i)}^{\pi/2} \int_0^{\bar{s}(q,p_i,\phi_0)} \sqrt{1 - r(q,p_i,\phi_0,\phi)}\, \mathrm{d}\phi\, \mathrm{d}\phi_0, & q > p_i. \end{cases} \tag{B.2}$$

Although the ϕ-integral can be given in terms of elliptic functions, this is best computed by evaluating the double integral numerically. The maximum value of $G_i(q)$ is obtained at $q = 1$:

$$G_i(1) = \int_0^{\pi/2} \int_0^{\phi_0} \mathrm{d}\phi\, \mathrm{d}\phi_0 = \frac{\pi^2}{8}.$$

Our basis functions G_i are closed-form expressions of those computed by Monte-Carlo sampling in Szabo & Kiss (2008). These can be used to determine the p-distribution, although we found the accuracy inferior to the solution based on the variation observable η, which was to expected.

In principle, we can expand G_i to $G_{ij}(q)$ for a (p_i,β_j)-grid in the same way that F_i were expanded to F_{ij}. However, a notable difference between the two-index basis functions of A- or q- data is that the $G_{ij}(q)$ all reach their maxima at the same point $q = 1$ since the two-point comparison can always contain two equal brightnesses for any p and β. Thus the β_j-curves of the $G_{ij}(q)$ of a given p_i form a curve family with the same abscissae for the minimum ($q = p_i$) and maximum ($q = 1$); that is, members of the family can easily be mimicked by a superposition of other members unlike in the case of $F_{ij}(A)$. A number of simulations indeed confirmed that G_{ij} are not usable for solving the inverse problem in practice; that is, β-information is not recoverable from q-data. Adding prior assumptions on the joint distribution did not help either, as it resulted in too heavy regularization, causing the solution to become almost entirely prior-based.

The same problem plagues the lightcurve slope estimate of McNeill et al. (2016) from two neighbouring points, exacerbated by the effect of the rotation period P. For any basis function, the abscissae for the two-point ratio s lie at some minimum $s_m(p, P)$ and maximum $s = 1$. The same abscissae apply not only to all β, but also to infinitely many other combinations of p and P that yield the same s_m. Thus any basis function can be mimicked by numerous different superpositions of basis functions at other p, P, and β, making the solution of the inverse problem ambiguous without heavy prior assumptions.

Appendix C: Stability properties of shape distribution from observed brightness variation

We can analyse the inverse problem of determining $f(p)$ with the same approach as in lightcurve inversion Kaasalainen & Lamberg (2006): we expand both the observed $C(A)$ and $f(p)$ as function series, and examine the relationship between their coefficients. This shows if all coefficients of $f(p)$ can be determined, and also how fast their errors grow as a function of their degree.

Let us expand $f(p)$ as the polynomial

$$f(p) = \sum_{n=1}^{\infty} c_n p^n; \quad p \in [0, 1].$$

For isotropic θ,

$$C(A) = \sum_n c_n \int_0^A p^n \frac{\sqrt{A^2 - p^2}}{\sqrt{1 - p^2}}\, \mathrm{d}p,$$

and from tables of integrals we find that this is

$$C(A) = A^2 \sum_n c_n A^n \frac{1}{n+1} F_1\left(\frac{n+1}{2}; \frac{1}{2}, -\frac{1}{2}; \frac{n+1}{2} + 1; A^2, 1\right),$$

where F_1 is the Appell hypergeometric function. This form can be transformed into the usual Gauss hypergeometric function ${}_2F_1$ so that

$$C(A) = A^2 \sum_n c_n A^n k_n G_n(A),$$

where

$$G_n(x) = {}_2F_1\left(\frac{n+1}{2}, -\frac{1}{2}; \frac{n+4}{2}; x^2\right) = \sum_j^{\infty} b_j^n x^{2j},$$

with

$$b_j^n = \frac{\left(\frac{n+1}{2}\right)_j \left(-\frac{1}{2}\right)_j}{j! \left(\frac{n+4}{2}\right)_j}; \quad (a)_j = \frac{\Gamma(a+j)}{\Gamma(a)}$$

(so $(a)_0 = 1 = b_0^n$), and

$$k_n = \frac{\sqrt{\pi}}{2(n+1)} \frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n+4}{2}\right)},$$

so $k_n \neq 0$ decreases monotonously as n increases, and $\lim_{n\to\infty} k_n = 0$. The decrease is moderate, approximated by, for example, $\sim(n+1)^{-1}[\log(n/2+3)]^{-3/2}$ for $n < 100$. For the gamma function, $\Gamma(n + 1/2) = \sqrt{\pi}(2n-1)!!/2^n$ and $\Gamma(n) = (n-1)!$.

Suppose the observed $C(A)$ is expanded (to hold for $0 \le A \le 1$) as

$$C(A) = A^2 \sum_{n=1}^{\infty} a_n A^n.$$

Then

$$a_1 = c_1 k_1 \Rightarrow c_1 = a_1/k_1; \quad c_2 = a_2/k_2;$$

$$a3 = c_3 k_3 + c_1 k_1 b_1^1 \Rightarrow c_3 = (a_3 - c_1 k_1 b_1^1)/k_3,$$

and so on recursively; that is,

$$c_n = \frac{1}{k_n}\left(a_n - \sum_{i=1}^{[n]-1} c_{n-2i} k_{n-2i} b_i^{n-2i}\right),$$

where $[n]$ is $(n+1)/2$ or $n/2$ for, respectively, odd or even n. Thus, all coefficients c_n are obtained, and their error grows as $1/k_n$, which is much slower than in, for example, lightcurve inversion.

We note that we can write a formal, more user-friendly one-to-one mapping between the polynomial coefficients determining $f(p)$ and $C(A)$. If $p \in [0, 1[$, and we expand (assuming $f(p)$ to vanish fast enough when $p \to 1$)

$$\frac{f(p)}{\sqrt{1-p^2}} = \sum_{n=1}^{\infty} d_n p^n,$$

we have

$$C(A) = \sum_n d_n \int_0^A p^n \sqrt{A^2 - p^2}\, \mathrm{d}p$$

$$= A^2 \sum_n d_n A^n \frac{{}_2F_1\left(-\frac{1}{2}, \frac{n+1}{2}; \frac{n+3}{2}; 1\right)}{n+1},$$

which is simply

$$C(A) = A^2 \sum_n d_n k_n A^n,$$

so we obtain a simple relationship between data and model:

$$d_n = \frac{a_n}{k_n}.$$

We also note that the above applies to the general triaxial ellipsoid as well. Let us now have a fixed $c \neq 1$, $b = 1$, and $a = 1/p$. Then

$$A^2 = \frac{p^2 \sin^2\theta + c^{-2}\cos^2\theta}{\sin^2\theta + c^{-2}\cos^2\theta},$$

so the iso-A curves are given by

$$\cos^2\theta_{A3}(p) := g_{A3}(p) = \frac{A^2 - p^2}{h(A) - p^2},$$

where

$$h(A) := A^2(1 - c^{-2}) + c^{-2}.$$

Now

$$C(A) = \sum_n c_n \int_0^A p^n \sqrt{g_{A3}(p)}\,\mathrm{d}p,$$

and this is

$$C(A) = \frac{A^2}{\sqrt{h(A)}} \sum_n c_n A^n \frac{1}{n+1} F_1\left(\frac{n+1}{2}; \frac{1}{2}, -\frac{1}{2}; \frac{n+1}{2} + 1; \frac{A^2}{h(A)}, 1\right).$$

This can be used to define a series expansion for the observed $C(A)$ with new basis functions instead of polynomials, so we have the same kind of one-to-one correspondence as above.

Publication III

Nortunen, H. and Kaasalainen, M., "LEADER: fast estimates of asteroid shape elongation and spin latitude distributions from scarce photometry," *Astronomy and Astrophysics*.

A&A 608, A91 (2017)
DOI: 10.1051/0004-6361/201731360

Astronomy & Astrophysics

LEADER: fast estimates of asteroid shape elongation and spin latitude distributions from scarce photometry

H. Nortunen and M. Kaasalainen

Tampere University of Technology, Department of Mathematics, PO Box 553, 33101 Tampere, Finland

Received 13 June 2017 / Accepted 25 September 2017

ABSTRACT

Context. Many asteroid databases with lightcurve brightness measurements (e.g. WISE, Pan-STARRS1) contain enormous amounts of data for asteroid shape and spin modelling. While lightcurve inversion is not plausible for individual targets with scarce data, it is possible for large populations with thousands of asteroids, where the distributions of the shape and spin characteristics of the populations are obtainable.
Aims. We aim to introduce a software implementation of a method that computes the joint shape elongation p and spin latitude β distributions for a population, with the brightness observations given in an asteroid database. Other main goals are to include a method for performing validity checks of the algorithm, and a tool for a statistical comparison of populations.
Methods. The LEADER software package read the brightness measurement data for a user-defined subpopulation from a given database. The observations were used to compute estimates of the brightness variations of the population members. A cumulative distribution function (CDF) was constructed of these estimates. A superposition of known analytical basis functions yielded this CDF as a function of the (shape, spin) distribution. The joint distribution can be reconstructed by solving a linear constrained inverse problem. To test the validity of the method, the algorithm can be run with synthetic asteroid models, where the shape and spin characteristics are known, and by using the geometries taken from the examined database.
Results. LEADER is a fast and robust software package for solving shape and spin distributions for large populations. There are major differences in the quality and coverage of measurements depending on the database used, so synthetic simulations are always necessary before a database can be reliably used. We show examples of differences in the results when switching to another database.

Key words. methods: analytical – methods: numerical – methods: statistical – minor planets, asteroids: general – techniques: photometric

1. Introduction

Photometric observations of total (disk-integrated) brightnesses are by far the most abundant source of information on asteroids in the sense of population coverage (Ďurech et al. 2015). There are currently some thousand shape and spin models from photometry for individual asteroids – for example, see the Database of Asteroid Models from Inversion Techniques (DAMIT[1]) site – and tens of thousands more will be obtained from sparse photometry from various sky surveys such as Panoramic Survey Telescope and Rapid Response System (Pan-STARRS), Gaia, and Large Synoptic Survey Telescope (LSST). Infrared photometry from, for example, the Wide-field Infrared Survey Explorer (WISE) mission can also be added to visual data to obtain size and surface material parameters for thousands of targets (Ďurech et al. 2016).

Even larger population-level attributes can be obtained by using all available photometric measurements from the rich survey databases also when the data are not sufficient for individual models, as shown in Nortunen et al. (2017) and Cibulková et al. (2017). We refer to such data as scarce photometry to distinguish it from sparse photometry. The possibility of obtaining population-level attributes is based on the principle of using even a few data points per target to construct a set of brightness variation estimates for a given population (defined by the user). As shown in Nortunen et al. (2017), such estimates are sufficient for robustly reconstructing the shape elongation and (ecliptically symmetric) spin latitude distributions of the population when there are thousands of samples available. The determination of the spin distribution requires the population to be dominated by orbits concentrated near the ecliptic plane.

In this paper, we describe in detail the software package Latitudes and Elongations of Asteroid Distributions Estimated Rapidly (LEADER) for obtaining the shape and spin distributions. The mathematical principles are presented in Nortunen et al. (2017), but the computational aspects require further exposition given here. The main issues are the choice of the grid discretization level in solving the inverse problem, experimenting with various setup choices to establish the stability of the result, the "deconvolution" of the result distributions in the shape-spin plot plane, and the mandatory use of simulations (from synthetic data for the same observing geometries as in the input database) to check the reliability of the inversion. We emphasize that, while the method itself is quite robust in the sense that the inaccuracy of the underlying ellipsoidal model is well tolerated since only the large-scale elongation and spin estimates are obtained, the properties and coverage of the database entirely dictate the reliability of the results.

This paper is organized in the following manner. In Sect. 2 we present the main algorithm, some computational details, and the visualization by deconvolution, while in Sect. 3 we discuss the implementation of the necessary simulations from synthetic

[1] http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php

data and the comparison routines between populations. In Sect. 4 we illustrate the use of the LEADER package with examples. We sum up in Sect. 5, and in the appendix we describe some main components of the software.

2. Main algorithm for computing distributions

The main algorithm consisted of roughly three phases. First, we had a forward model for computing the brightness variation estimate η from observations. Then, we considered the inverse problem and determined the joint (p, β) distribution. Finally, we plotted the results and applied a deconvolution filter to smoothen the solution.

2.1. Forward model: brightness variations

We considered a population of N asteroids. For our analysis, we required a large population, with $N \gtrsim 1000$. In the forward model, our observable is the brightness deviation estimate η. Our objective is to obtain one or more ηs for each asteroid. With the observed brightness values L, we defined η as

$$\eta = \frac{\Delta(L^2)}{\langle L^2 \rangle}, \tag{1}$$

where $\Delta(L^2)$ is a measure of variation for L^2 as defined in Nortunen et al. (2017):

$$\Delta(L^2) = \sqrt{\langle (L^2 - \langle L^2 \rangle)^2 \rangle}.$$

We used the squared brightness L^2 for convenience; a more detailed explanation is given in Nortunen et al. (2017). We were then able to compute the amplitude A from each η, and from all the amplitude values of the population, we constructed the cumulative distribution function (CDF) $C(A)$. To obtain the brightness deviation estimates, we analyzed the data file of asteroid i, where $i = 1, \ldots, N$, repeating the following steps for each file:

1. From the data file, we read the Julian date, intensity, and the directions of the Earth and Sun (these should be computed if not given) as seen from the asteroid translated into the origin. We denoted the direction vectors of the Earth and Sun $\boldsymbol{e}_{\text{earth}}$ and $\boldsymbol{e}_{\text{sun}}$, respectively. For each data point, we computed the phase angle α between the Sun and the Earth, and required that

   $$\alpha := \arccos(\boldsymbol{e}_{\text{sun}} \cdot \boldsymbol{e}_{\text{earth}}) \leq \alpha_{\text{tol}}. \tag{2}$$

   In Nortunen et al. (2017), we used a limit of $\alpha_{\text{tol}} = 30°$ for the WISE database. For the Pan-STARRS1 database, we used $\alpha_{\text{tol}} = 20°$. We filtered out each brightness measurement where the phase angle exceeds this given tolerance. In addition, we required that a set of measurements has been done within a small enough change in geometry, and at least five brightness values are required for a valid η estimate. For Pan-STARRS1, the phase angle α changes rapidly in time, so we could have used a condition that all measurements in a set are done within, for example, three days, to keep the change in the phase angle small. For the WISE database, more liberal rules can be used with the time span, as the phase angle changes at most 0.4° within a one-week-long set of measurements. With this set of rules, we separated the measurements on a single data file into multiple sets. After that, we analyzed each set individually for computing η.
2. *Phase angle correction.* Depending on the phase angle α, we should apply either an exponential or a linear correction to the brightness values $L(\alpha)$, as explained in Kaasalainen et al. (2001). Let us consider a single set of measurements. If $\min_i \alpha_i < 8°$ in the set, we required that the phase angle changes no more than one degree, that is:

   $$(\max_i \alpha_i) - (\min_i \alpha_i) \leq 1°. \tag{3}$$

   If the phase angle does change for more than one degree, we applied an exponential correction. In other words, we applied an exponential curve $a \exp(b\alpha)$ into the (α, L) data, and used it to normalize the brightness data into the form

   $$\frac{L(\alpha)}{a \exp(b\alpha)}, \quad a > 0,\ b < 0. \tag{4}$$

   If $\min_i \alpha_i \geq 8°$ in the set, we required that the phase angle changes at most two degrees:

   $$(\max_i \alpha_i) - (\min_i \alpha_i) \leq 2°. \tag{5}$$

   If the phase angle changes more, we applied a linear correction by using a least squares fit of the form $(a\alpha + b)$ into the (α, L) data, and use it to normalize the brightness data into

   $$\frac{L(\alpha)}{a\alpha + b}, \quad a < 0. \tag{6}$$

   We note that with a small number of points, the correction may be unstable and can be omitted.
3. For each measurement set, we computed the brightness variation η using Eq. (1). In Nortunen et al. (2017), we derived how η and the amplitude A are directly related:

   $$A = \sqrt{1 - \left(\frac{1}{\sqrt{8}\eta} + \frac{1}{2} \right)^{-1}}. \tag{7}$$

   We omitted any complex-valued or non-finite amplitudes.

When all the brightness variations and amplitudes (η and A) have been computed, we sorted the amplitudes in an increasing order. Then the CDF of A is simply $C(A_i) = i/k$ for $i = 1, \ldots, k$.

2.2. Inverse problem: obtaining the joint distribution for latitudes and elongations

In the inverse problem, our objective is to determine the distributions of two parameters, the shape elongation $p \in [0, 1]$ and the spin latitude $\beta \in [0, \pi/2]$. We modelled the asteroids with the shape of a triaxial ellipsoid, with semiaxes a, b and c, and $a \geq b = c = 1$. With this model, the shape elongation is simply $p = b/a$. Here a small p value corresponds to an elongated body, while $p = 1$ describes a sphere. For the spin latitude, $\beta = 0$ represents a spin direction that is perpendicular to the ecliptic plane, while $\beta = \pi/2$ means the spin direction is in the ecliptic plane[2]. Our model does not have any way of distinguishing whether the spin latitude is above or below the ecliptic plane. We assumed that the observations are concentrated near the ecliptic plane. This assumption does not usually hold entirely; the amount of variation in the ecliptic latitudes of the observations depends on

[2] We note that in our convention, β is the complementary angle of the traditionally used β. This is due to mathematical purposes. To avoid confusion, we use radians instead of degrees for the values of our β in plots.

the orbits of the population and the sampling epochs in database used. As a result, the computed β distribution tends to be less accurate than the p distribution. We tested the validity of the assumption, and for the databases we used in this paper, some 95% of the observations were concentrated within a $\pm 20^\circ$ sector near the ecliptic plane.

To solve the inverse problem, we created a grid of $N_p \times N_\beta$ bins for our (p,β) values. For our algorithm, we typically used $N_p = 20$ and $N_\beta = 29$, so every bin is approximately 0.05×0.05 units in size, and equally spaced. We chose (p_i,β_j) as a random point near the centre of each bin. Alternatively, as the values $p < 0.4$ are expected to have lower occupation numbers than the higher p values, we may lower the resolution for such values. Similarly, since the occupation numbers for β are expected to be somewhat proportional to $\sin\beta$ (this means a uniform density on a sphere), low β values were expected to have lower occupation numbers, so we may lower the resolution of, for example, β values smaller than $\pi/4$.

In Nortunen et al. (2017), we derived how the CDF $C(A)$ can be expressed in an analytical integral form when we have infinite observations available in every geometry. When we used a grid of points (p_i,β_j), we were able to express $C(A)$ as the following superposition:

$$C(A) = \sum_{ij} w_{ij} F_{ij}(A). \tag{8}$$

Here w_{ij} are the occupation numbers (weights) of each bin (p_i,β_j), and

$$F_{ij}(A) = \begin{cases} 0, & A \le p_i \\ \frac{\pi}{2} - \arccos \frac{\sqrt{A^2-p_i^2}}{\sin\beta_j \sqrt{1-p_i^2}}, & p_i < A < \mathcal{F}(p_i,\beta_j) \\ \frac{\pi}{2}, & A \ge \mathcal{F}(p_i,\beta_j) \end{cases} \tag{9}$$

are analytical basis functions, where

$$\mathcal{F}(p_i,\beta_j) = \sqrt{\sin^2\beta_j + p_i^2 \cos^2\beta_j}.$$

Next, we constructed the data matrix M (with k rows and $N_p \cdot N_\beta$ columns) such that each column of M contains a basis function $F_{ij}(A)$. When we write $C(A) =: C \in \mathbb{R}^k$, the superposition of Eq. (8) can be written as a linear system,

$$Mw = C, \tag{10}$$

where the unknown vector $w \in \mathbb{R}^{N_p \cdot N_\beta}$ contains the occupation numbers w_{ij} of each bin. Before solving the system, the use of regularization is highly recommended, especially for the more unstable β. Let R_p be an $((N_p - 1) \cdot N_\beta) \times (N_p \cdot N_\beta)$ matrix that is meant to smooth the solution for p, and R_β be the respective $(N_p \cdot (N_\beta - 1)) \times (N_p \cdot N_\beta)$ regularization matrix for β. For indices ij, we have:

$$(R_p)_{ij} = \begin{cases} -1/(p_{i+1} - p_i), & i = j \\ 1/(p_{i+1} - p_i), & j = i+1 \\ 0, & \text{elsewhere,} \end{cases}$$

and similarly for R_β. The regularization matrices approximate the gradients at each w_{ij} in the p- and β-directions. For the regularization parameters, we typically used values $\delta_p = 0.1$ and $\delta_\beta = 1$. Now, let us create an extended matrix $\tilde{M}$ and an extended vector $\tilde{C}$:

$$\tilde{M} = \begin{pmatrix} M \\ \sqrt{\delta_p} R_p \\ \sqrt{\delta_\beta} R_\beta \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} C \\ 0_{(N_p-1)N_\beta} \\ 0_{N_p(N_\beta-1)} \end{pmatrix},$$

with our extended linear system being

$$\tilde{M}w = \tilde{C}. \tag{11}$$

To obtain the occupation numbers w_{ij}, we solved for w from Eq. (11) by using, for example, MATLAB's[3] linear least squares method with a positivity constraint $w_{ij} \ge 0$. The peak of the joint (p,β) distribution is simply the (p_i,β_j) bin with the highest occupation number w_{ij}.

2.3. Visualization

To estimate the goodness of the fit $C(A) = \sum_{ij} w_{ij} F_{ij}(A)$ from Eq. (8), or the equivalent form $C = Mw$ from Eq. (10), we may plot C and Mw in the same plot. The relative error $\|C - Mw\| / \|C\|$ is usually less than 1% when the population contains at least 1000–2000 objects. For the actual joint distribution $f(p,\beta)$, where $f(p_i,\beta_j) = \tilde{w}_{ij}$ and $\tilde{w}_{ij}$ are the occupation numbers w_{ij} normalized such that $\sum_{i,j} f(p_i,\beta_j) = 1$, we may plot the solution $(p,\beta,f(p,\beta))$ either as a three-dimensional surface plot, or alternatively as a contour plot. The marginal density functions (DFs) can also be computed for p and β:

$$f(p_i) = \sum_{j=1}^{N_\beta} \tilde{w}_{ij}, \quad f(\beta_j) = \sum_{i=1}^{N_p} \tilde{w}_{ij}. \tag{12}$$

As a post-processing tool, we may apply deconvolution to correct for dispersion in the obtained solution. The deconvolution is used as a primarily visual tool, and it is applied only for the joint (p,β) distribution, not the marginal DFs. In order to know what kind of post-solution correction is needed, synthetic simulations (Sect. 3.1) should be performed on the database used. With the synthetic simulations, we gained understanding of the accuracy levels of the method, and were able to detect systematic errors associated with the database. As the solution tends to spread when moving away from the peak, it is a common procedure to introduce dampening to bins away from the peak. Let i^* and j^* be the indices for the statistical peak of the solution, that is, the bin with the highest occupation number. Then, the dampening we applied is

$$\tilde{w}_{ij}^{\text{corr}} = \frac{\tilde{w}_{ij}}{(|i^* - i| + |j^* - j| + 1)^n}. \tag{13}$$

Usually it suffices to choose $n = 1$, but if heavier dampening is required, we may choose a larger n. As the solution of the shape elongation p is often shifted too much to the left (towards more elongated values), for example, by the amount $\Delta\mathcal{P} \ge 0$, we may additionally shift the p values to the right (towards more spherical values):

$$p_i^{\text{corr}} = \min(p_i + \Delta\mathcal{P}, 1). \tag{14}$$

For the WISE database, we chose $\Delta\mathcal{P} = 0.1$ in Nortunen et al. (2017). The more noisy the database is, the higher the required shift $\Delta\mathcal{P}$ is. For β, the error behaviour is much harder to model and may lead to exaggerated correction, so typically we did not apply any correction in the β direction. However, we acknowledge that the solution tends to avoid extreme ends, so values near $\beta = 0$ (perpendicular to the ecliptic plane) and values near $\beta = \pi/2$ (in the ecliptic plane) have a tendency to shift away from the end points, moving towards the middle.

The computation times of the main algorithm are negligible. Depending on the size of the inspected population, reading

[3] The MathWorks, Inc, MATLAB R2014a, Natick, 2014.

the geometries from a database may take a few minutes, while the computation of the solution via the inverse problem is even faster. Therefore, it is easy to experiment with different grids for the solution of the inverse problem, or to test different populations. The latter means it is also fast to compare populations, which we will discuss in more detail in Sect. 3.2.

3. Main implementations of LEADER

In Sect. 3.1, we discuss an essential test when using the LEADER package: accuracy estimation by running simulations on synthetic data. Simulations are the only way to gain an understanding of how applicable the main algorithm is for a given database, and they should be performed on every database before those can be reliably used. In Sect. 3.2 we describe an extension of the main algorithm, an application for comparing shape elongation and spin latitude distributions of two populations.

3.1. Verifying the method using simulations based on synthetic data

The only way to test the correctness of the obtained solution is to run simulations based on synthetic data, where the $(p,\ \beta)$ distribution of the artificial population is known, and see how accurately the solution is obtained. The test can additionally be used to detect systematic errors. The level of accuracy has a strong dependence on the database used (Nortunen et al. 2017), so whenever we start to use a new database, it is necessary to run simulations to see how well our method performs with the database.

The synthetic simulations begin by choosing a single peak for the $(p,\ \beta)$ distribution. Let us denote this peak (p^*,β^*). Now, let us assume we run the simulation for N asteroids in a population. Then, we repeat the following steps for each asteroid i, where $i = 1, \ldots, N$:

1. We choose an asteroid model from DAMIT[4], with a shape elongation $p = b/a$ (here a is the longest diameter in the equatorial xy-plane, and b is the width in the corresponding orthogonal direction) that is close to the peak value p^*. For example, we have set a criterion,
$$|p - p^*| \leq 0.075. \tag{15}$$
We may apply basic transformations, such as stretching, on DAMIT objects in order to get the intended shape elongation value for the asteroid. We compute the normal and area for each facet of the body.
2. The next step is to construct the brightness data and our η estimate. We choose a β from Gaussian distribution, with β^* as the mean value, and 0.05 as the standard deviation (with a restriction that $\beta \in (0,\ \pi/2)$). We also fix the longitude λ by choosing it from a uniform distribution, $\lambda \in [0, 2\pi]$.
3. Next, we need geometries from the asteroid database we are studying. We read the data from a data file belonging to that database. We extract the direction vectors of the Sun and the Earth from the data file, and filter out the cases when the condition of Eq. (2) is violated. Next, we transform these vectors to the asteroid-fixed frame using a coordinate transformation (Kaasalainen et al. 2001), with our fixed β and λ. We denote the direction vectors in the asteroid's own frame $\boldsymbol{e}_{\rm sun}$ and $\boldsymbol{e}_{\rm earth}$. Then we compute the total brightness L for the target. Finally, we add a minor Gaussian perturbation to L to simulate noise.
4. From now on, we proceed as in Sect. 2.1. We require at least five L values for a valid η estimate, and if necessary, we only consider the measurements that have been done within a short time span to keep the changes in geometry small. If necessary, we apply phase correction on the sets, and finally, we compute η and A for each set as in Eqs. (1) and (7). We sort the amplitudes in an increasing order, and construct the CDF $C(A)$ as in Sect. 2.1. As the actual p and β distributions are known with the synthetic data, it is often illustrative to plot the marginal DFs and the contour plot of the joint $(p,\ \beta)$ distribution.

The next phase is to obtain the solution of the inverse problem. This is done in the same way as that described in Sect. 2.2. The graphical presentation of the results, as well as possibly applying a deconvolution filter, is done identically to Sect. 2.3.

3.2. A comparison of two populations

A comparison between user-determined populations is typically desirable. Taking into account that the database used tends to cause biases in the "absolute" values of distributions, the relative differences between distributions can be expected to be more robust results from the database.

For populations S_1 and S_2, we ran the algorithm described in Sect. 2. We collected the p, $f(p)$, β and $f(\beta)$ information from both algorithms, where p and β are the grid points, and $f(p)$ and $f(\beta)$ are the marginal density functions, computed from the normalized occupation numbers $\tilde{w}_{ij}$ as in Eq. (12).

First, we computed the CDFs for the marginal DFs, denoted F_p and F_β:
$$F_{p_m} = \sum_{i=1}^{m} f(p_i), \quad F_{\beta_n} = \sum_{j=1}^{n} f(\beta_j). \tag{16}$$
With the CDFs $F_p(S_1)$, $F_p(S_2)$, $F_\beta(S_1)$ and $F_\beta(S_2)$ computed, we were able to compute the statistical differences as defined in Nortunen et al. (2017):
$$\begin{cases} D_p(S_1, S_2) = \alpha_k \left\| F_p(S_1) - F_p(S_2) \right\|_k \\ D_\beta(S_1, S_2) = \alpha_k \left\| F_\beta(S_1) - F_\beta(S_2) \right\|_k , \end{cases} \tag{17}$$
where we computed cases $k = 1$, $k = 2$ and $k = \infty$, with the scaling factors $\alpha_1 = 1/4$, $\alpha_2 = 1$ and $\alpha_\infty = 2$. As a general rule of thumb, the statistical difference between two populations can be considered significant if $D \gtrsim 0.2$. Naturally, one number does not tell everything about the quality of the statistical difference, which is why we used several different norms. Plotting the DFs (and CDFs) of both populations in the same figure is often more illustrative in terms of analyzing differences.

4. Tests with synthetic data and other examples

In this section, we demonstrate two applied examples of our method. First, we run a series of synthetic simulations to test the validity of the WISE database. In all of our examples, we used a customized version of the WISE database, which was compiled from the original data in Ďurech et al. (2016). The algorithm described in Sect. 3.1 was executed several times for different $(p,\ \beta)$ peak values and population sizes. In the second example, we computed the $(p,\ \beta)$ distributions for two populations using the algorithm from Sect. 2, and compared the distributions using the algorithm from Sect. 3.2. Both examples were executed using MATLAB software[3].

[4] We have chosen DAMIT as the source for synthetic models doe to its rich variety of realistic shapes.

We checked the accuracy of our method by observing how well we were able to reconstruct the peak of the joint (p, β) distribution. In our setup, we took a population of N asteroids from DAMIT, with the shape elongation and spin latitude and longitude known for each asteroid model. We used the geometries of the asteroid, the Sun, and the Earth that were computed for the WISE database in Ďurech et al. (2016). We observed how accurately the (p, β) peak was computed for the population, and repeated this 50 times, each time having a different, randomly generated (p, β) peak. This way, we gained a good understanding of how accurately the method computes the peaks for the given population size, no matter where the most occupied bin lies in the (p, β)-plane. After this, we repeated the same simulation setup for another population size; we considered populations ranging from 100 to 5000 asteroids in order to see how the accuracy of the method improves with a growing number of asteroids in the population.

Plots of different population sizes are presented in Figs. 1 and 2. Each plot draws the actual p or β peak versus the computed p or β peak, respectively, also showing the ideal case when the solution is completely accurate. As we can see from the plots in Fig. 1, the variance in the p peak decreases noticeably when the population size increases. In addition, there is an obvious systematic error that the p peak has been shifted "down", towards more elongated shapes. This shift is mainly due to the model (and data) noise. For example, a spheroidal shape with surface irregularities, estimated at $p = 1$, produces photometric variation interpreted as $p < 1$ by the smooth ellipsoidal model. For peaks with $p \gtrsim 0.6$, the shift is about 0.1 units, whereas for lower peaks, there is a bigger shift. In realistic populations, shapes with $p < 0.5$ are rare, so it is safe to assume that the value of the p peak is much higher than 0.5. Hence, it usually suffices to expect that the computed p peak is 0.1 units too low for WISE data.

In order to check whether the above result is a bias related to the asteroid database or our way of determining the shape elongation, we repeated the simulations with an alternative way to compute p. We considered the contours of the DAMIT-based shapes and computed the ellipsoid (with semiaxes a and b, where $a \geq b$) that best fits the contours, and then we computed the shape elongation: $p = b/a$. We observed that this produced no difference in the simulations, so we conclude that the bias is caused by the database; the bias caused by our definition of p is random rather than systematic. From now on, we will use our original definition of p (rather than finding the best elliptical fit for contours), as it is more simple and computationally faster. To correct the systematic error encountered with the WISE database, we shift the p values up by 0.1 units in the deconvolution phase.

As we can see in Fig. 2, the location of the computed β peak is a coarse approximation of the actual position. The variance in the accuracy is large. While a bigger population does improve the accuracy of the peak, the variance always remains to some extent. Low β values perpendicular to the ecliptic plane have a lower variance, while the high β peaks in the ecliptic plane have a high variance, suggesting that the solution is moderately unstable if the actual β is high. Since the tail of the computed distribution tends to spread towards the ecliptic plane, as seen in Nortunen et al. (2017), it is not surprising that the computed peak tends to shift away from the ecliptic plane. Similarly for low β values, the computed peak shifts away from the low end of the β range. A systematic correction for β is complicated to implement, so we merely note that caution should be used with the obtained β solution, as our method yields coarse estimates of the spin latitude distribution.

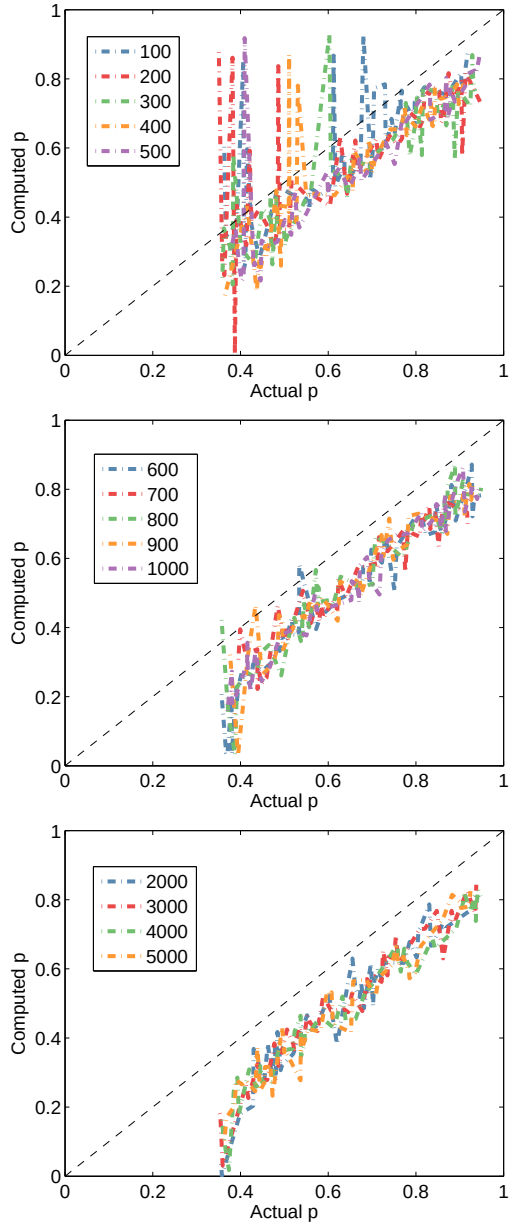

Fig. 1. Synthetic simulations illustrating how the accuracy of the p solution increases when the population size increases from 100 to 5000 asteroids, with the geometries from WISE. The black, dashed "$y = x$" line presents the ideal case when the computed solution is completely accurate.

The simulations presented in Figs. 1 and 2 were done with an equally spaced (p, β) grid. We tried simulations where we utilized a lower resolution for different p and β values, but we noticed that it reduced the accuracy of the method, especially for β, when using WISE database. The most accurate results were obtained by using an equally spaced grid. Alternatively, as it is extremely unlikely that $p < 0.4$, we could cut our grid by including only values of $p \in [0.4, 1]$. However, we choose to include

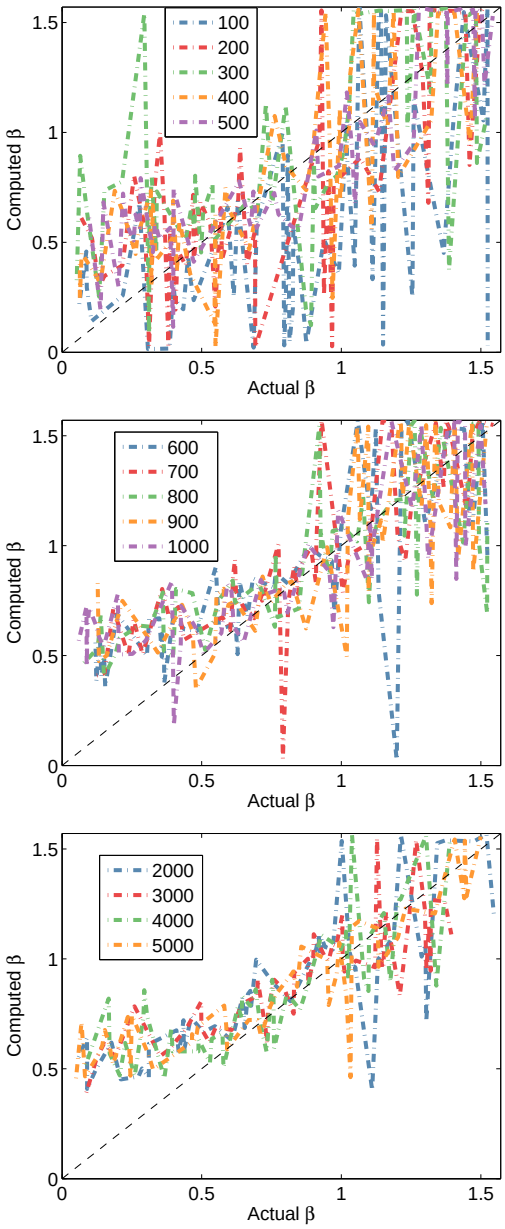

Fig. 2. Synthetic simulations similar to the ones in Fig. 1, but for β.

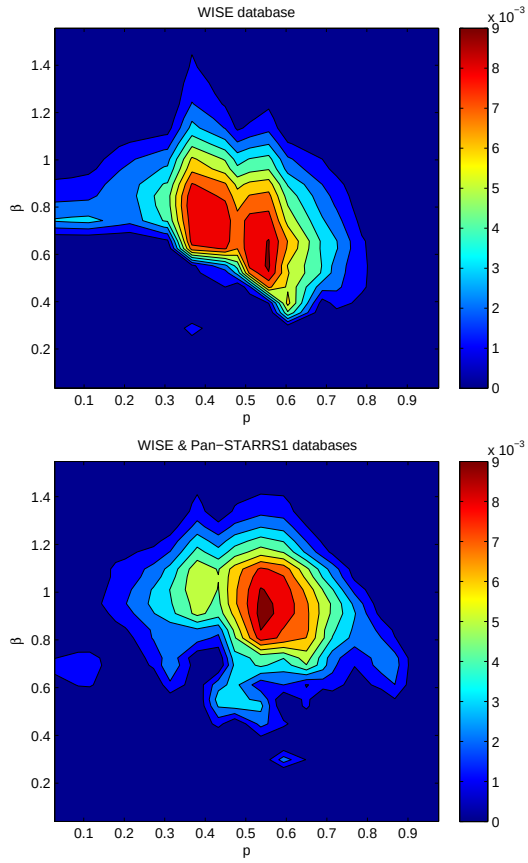

Fig. 3. Contour solution of the joint (p, β) distribution, computed from WISE (*top*) and combined WISE & Pan-STARRS1 (*bottom*) databases.

the whole interval [0, 1] for completeness, since having high occupation numbers on low p values is usually a good indicator of some systematic error; that is, low p values in the grid are useful for error checking.

For future research, we are interested in testing our method with the Pan-STARRS1 database (Cibulková et al. 2017). We tested the accuracy of the method with Pan-STARRS1 using synthetic simulations, and found that the shape elongation p is highly accurate, whereas the solution of the spin latitude β is more unstable, as can be expected due to the scatter in the observation geometries. A more detailed analysis of using our method with the Pan-STARRS1 database is presented in Cibulková et al. (2017). We are also interested in whether we can "combine" the databases by taking η estimates computed from measurements taken from both WISE and Pan-STARRS1 databases. Therefore, we have performed a preliminary examination on whether our method works and is accurate with combined databases. When we are trying to determine the shape and spin distributions of a certain subpopulation, such as an asteroid family, we may not always have a sufficient number of targets in separate asteroid databases. In such cases, we may attempt to supplement the number of ηs by taking observations from multiple databases, provided our method remains accurate.

To estimate the bias caused by databases, we inverted a subpopulation of about 70 000 asteroids from the WISE database, and compared the results with those obtained from a joint WISE and Pan-STARRS1 sample (containing the aforementioned WISE asteroids, and an additional subset of about 70 000 asteroids from the Pan-STARRS1 database that is used in Cibulková et al. 2017). In the latter sample, we took the brightness variations η from each asteroid, and used the combined ηs to construct the CDF $C(A)$. Here we only considered the asteroids that had brightness data available in both databases; this was to avoid selection effects, that is, to ensure that the differences between databases are not caused by observed targets being different. We plotted the computed (p, β) distributions in Fig. 3, and a

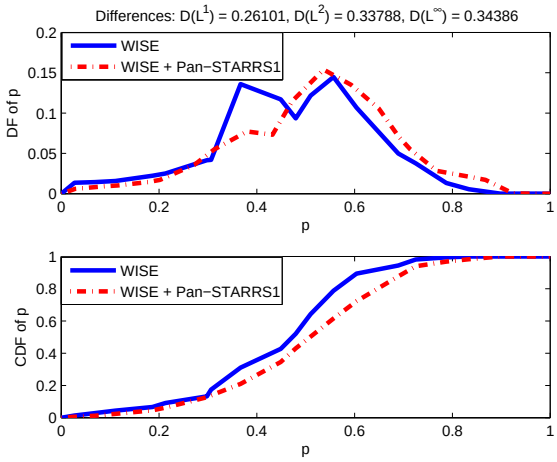

Fig. 4. Comparison of the marginal DFs (*top*) for WISE population and combined WISE & Pan-STARRS1 databases, and of their marginal CDFs (*bottom*), for shape elongation p.

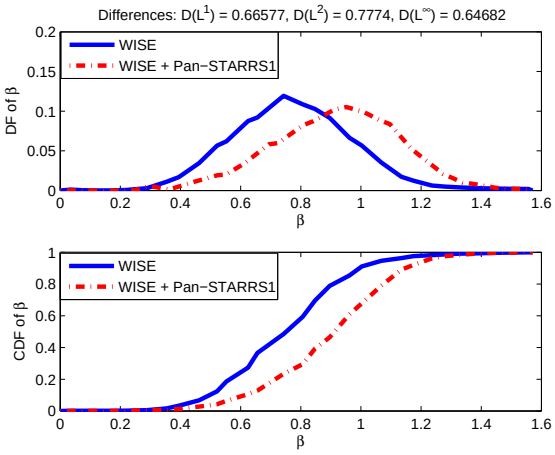

Fig. 5. Comparison of the marginal DFs (*top*) for WISE population and combined WISE & Pan-STARRS1 databases, and of their marginal CDFs (*bottom*), for spin latitude β.

comparison of the marginal DFs and their CDFs in Fig. 4 for p, and in Fig. 5 for β. No deconvolution has been used for the joint distributions of Fig. 3, as we wanted to preserve the information about multiple peaks as well as the spreading behaviour.

Even with half of the population taken from WISE, the combined WISE and Pan-STARRS1 population provides noticeably different distributions compared to merely considering the WISE population. For the marginal distributions of the shape elongation, the differences are minor, despite the moderately high $D(L^i)$ values in Fig. 4. The second double peak is so dominant in the combined database that it dampens the more elongated peak observed from the mere WISE population. In addition, the right tail of the distribution is somewhat heavy when using combined databases. Nevertheless, a visual inspection indicates that these differences in the p distributions are small, and should not be considered significant. Meanwhile, the differences are greater in the marginal distributions of the spin latitude in Fig. 5, as the β peak has been shifted towards the ecliptic plane in the combined population.

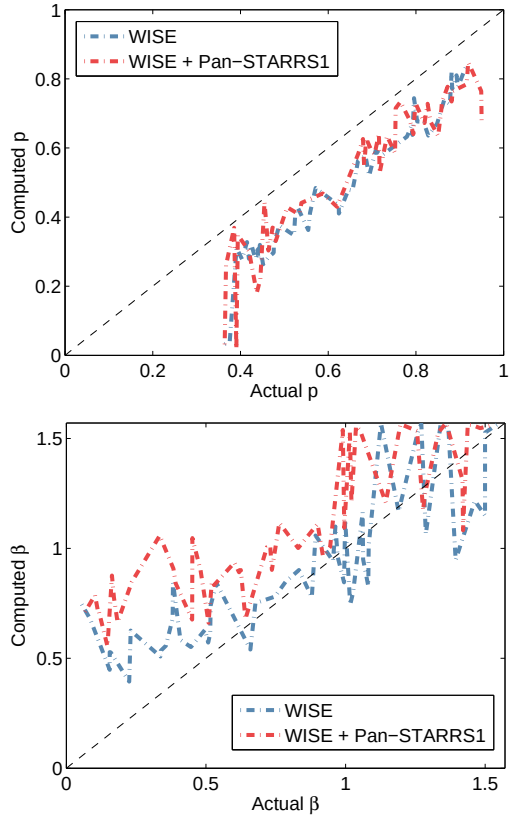

Fig. 6. Synthetic simulations illustrating the accuracy of the p (*top*) and β (*bottom*) solutions, using geometries from WISE and a combined WISE & Pan-STARRS1 database. The black, dashed "$y = x$" lines present the ideal case.

To determine whether the differences are caused by our method or the databases, we ran synthetic simulations, using geometries obtained from the above-mentioned databases. Once again, we only took geometries from the targets that had been in observed in both databases. The results from the synthetic simulations have been plotted in Fig. 6, in a format similar to that in Figs. 1 and 2. The accuracy of the p solution is clearly unaffected by the combining of the databases, so the slight differences in the computed p distributions in Fig. 4 are likely to be caused by the differences in the databases. As β is more sensitive to the observing geometries, the accuracy of the β solution deteriorates faster than that of the p solution when we combine databases.

We conclude that combining databases may lead to instabilities, so combining databases may not be useful if they are constructed in a different manner. If a database contains enough targets in a subpopulation, combining it with other databases will not necessarily increase the accuracy of the method. Furthermore, this example shows how much bias the choice of database introduces, which further emphasizes that no database should be blindly used with our method. Instead, every time we are introduced to a new database, it is crucial to run synthetic simulations in order to test the validity and error levels of the method.

5. Discussion and conclusions

We have introduced the LEADER software package for the fast estimation of population-level distributions of shape elongations and spin latitudes of asteroids. The method itself is demonstrably robust and designed to extract all the available information from databases that contain at least a few photometric points for each target (and preferably thousands of targets), when the data are not sufficient for individual models of the targets. However, we cannot overemphasize the necessity of testing the usefulness of the database with simulations based on synthetic data at the same geometries as in the database. This is the only way to assess the reliability of the inversion result, which is essentially dictated by the database.

The synthetic simulations provide a good overview of the applicability of the database. If the geometries are undersampled or the number of reasonably good brightness variation estimates is low, the acquisition of useful population-level distribution estimates is not possible regardless of the method. We have also shown how different the results become when switching to another database. For example, when combining the η estimates from both WISE and Pan-STARRS1, the bias introduced to the spin latitude is significant compared to using observations from merely the WISE database. As our examples show, some databases are more informative than others even if they all seem to be extensive at a first glance. The simulations are necessary to determine this.

Acknowledgements. We would like to thank Matti Viikinkoski for his valuable comments and feedback with the software implementation. This research was supported by the Academy of Finland (Centre of Excellence in Inverse Problems). H.N. was supported by the grant of Jenny and Antti Wihuri Foundation. This publication also makes use of data products from NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the Planetary Science Division of the National Aeronautics and Space Administration. In addition, this research made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory/California Institute of Technology, under contract with the National Aeronautics and Space Administration. We would also like to thank the reviewer for insightful comments that improved this paper. Last but not the least, we would like to thank Josef Ďurech and his research group for providing important feedback concerning the software package.

References

Cibulková, H., Nortunen, H., Ďurech, J., & Kaasalainen, M. 2017, A&A, submitted
Ďurech, J., Carry, B., Delbo, M., Kaasalainen M., & Viikinkoski, M. 2015, Asteroids IV, ed. P. Michel et al. (Tucson: U. Arizona), 183
Ďurech, J., Hanuš, J., Ali-Lagoa, V., Delbo, M., & Oszkiewicz, D. 2016, Proceedings of the International Astronomical Union, S318, 170
Kaasalainen, M., Torppa, J., & Muinonen, K. 2001, Icarus, 153, 37
Nortunen, H., Kaasalainen, M., Ďurech, J., et al. 2017, A&A, 601, A139

Appendix A: Structure of the software package

In this appendix, we present the overall structure of LEADER as a rough-scale pseudocode, as well as listing the key functions of the software package. First we discuss the main routine, then the variant used for synthetic simulations, and the application for comparing distributions. The software package presented here is designed for analyzing the WISE database.

The database is available for download in DAMIT, under the Software section[5].

A.1. Main routine

The principle of the main function `leader_main_WISE` has been presented on Table A.1. We assumed the population consists of N asteroids, and for each asteroid, we have a datafile available. For each datafile, we read the geometries, measurement times and brightness intensities using the function `lcg_read_WISE`. We split the observations into multiple sets according to the principles mentioned in Sect. 2.1: all measurements in the same set must be done within a small enough change in geometry, and at least five observations are required. Then for each set, we computed the brightness variation η and amplitude A using Eq. (7). After analyzing all datafiles, we had a list of amplitudes (the vector `A_vec` on Table A.1). We sorted the elements of the vector into an increasing order, and then we construct the CDF of A. In the inversion phase (function `leader_invert`), the matrix $\tilde{M}$ and the vector $\tilde{C}$ are generated as explained in Sect. 2.2. The MATLAB function `lsqnonneg` computes the solution to the non-negative least-squares problem

$$\min_{w} \left\| \tilde{M}w - \tilde{C} \right\|_2^2, \quad \text{where } w \geq 0.$$

Finally, we used the function `leader_plots` to plot the solutions, and the function `leader_postprocess_WISE` to deconvolute the solution visually.

A.2. Synthetic simulator

We have presented the principle of the synthetic simulator `leader_synth_main_WISE` on Table A.2. We assumed a population of N asteroids. For each target, we fixed a desired $(p_{\text{wanted}}, \beta_{\text{wanted}})$ value, and we chose the longitude λ_{wanted} from a random uniform distribution $[0, 2\pi]$. Then, we kept selecting randomized and stretched asteroid models from the DAMIT database, until we had an asteroid with the desired p value (within a certain tolerance). Then we read measurement dates and geometries (direction of the Sun and Earth) from the database we were testing (for example, WISE). We fixed $\beta = \beta_{\text{wanted}}$ and $\lambda = \lambda_{\text{wanted}}$, and used them to transform the direction vectors of the Sun and Earth into the asteroid's own frame. Then, for each geometry, we used a scattering law to compute the brightness intensity L, adding a small Gaussian noise. The rest of the algorithm is identical to the main routine: we split the data into measurement sets, computed η and A for each of them, constructed the CDF $C(A)$, used the subfunction `leader_invert` to compute the solution distribution, and finally visualized the solution with `leader_plots` and `leader_postprocess_WISE` subfunctions. It is recommended that some fine-tuning is done in the deconvolution function (`leader_postprocess_WISE`) to reshape the solution closer to the synthetic (p, β) distribution from the forward model.

[5] http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php?page=download_software

```
function leader_main_WISE

for i=1:N
   call function lcg_read_WISE(datafile i)
      read datafile
      split data into measurement sets
      compute eta for each set
      compute A from eta
      return A
   endfunction
end

A_sort = sort(A_vec)
for j=1:length(A_sort)
   CDFA(j) = j/length(A_sort)
end

call function leader_invert
   create (p, beta) grid
   create Mtilde, Ctilde
   W = lsqnonneg(Mtilde, Ctilde)
   return W
endfunction

call function leader_plots
   plot contour in (p, beta) plane
   compute marginal distributions
   draw marginal distributions for p and beta
endfunction

call function leader_postprocess_WISE
   find (p, beta) peak
   dampen bins away from the peak
   P_new = P + deltaP
   draw deconvoluted contour in (p, beta) plane
endfunction

endfunction
```

Table A.1. Pseudocode of the main function of LEADER.

A.3. Comparison of asteroid populations

The application for the statistical comparison of asteroid populations is called `ast_comparison_WISE`, and its principle has been presented on Table A.3. We ran the main routine `leader_main_WISE` first for population 1, saving the used p and β grids and their marginal DFs into variables `p1`, `beta1`, `fp1` and `fbeta1`. We did the same for population 2, saving the grids and their marginal DFs respectively into variables `p2`, `beta2`, `fp2` and `fbeta2`. Then we called the subfunction `KS_comparison`. The subfunction constructs the CDFs of each marginal DF, with `Cp1` and `Cp2` being the CDF of p for populations 1 and 2, and `Cb1` and `Cb2` being the CDF of β for populations 1 and 2, respectively. To compare the CDFs, we interpolated `Cp2` and `Cb2` at the grid points of population 1. Then we computed the statistical differences as defined by Eq. (17). Finally, we plotted the marginal DFs and their CDFs as in Figs. 4 and 5 to illustrate the differences of the distributions.

```
function leader_synth_main_WISE

for i=1:N
   set p_wanted, beta_wanted, lambda_wanted

   while ( |p-p_wanted| > tol )
      call function damit_model
         read random datafile
         return vertex and face information
      endfunction
      apply stretch on vertices
      call function leader_ellipsoid
         compute a, b, c, p = b/a
         return p
      endfunction
   endwhile

   call function
   leader_brightness_synth_WISE(datafile i)
      beta=beta_wanted, lambda=lambda_wanted
      call function lcg_read_synth_WISE
         read datafile
         return dates, geometries
      endfunction
      compute L for each geometry, add noise
      split data into measurement sets
      compute eta for each set
      compute A from eta
      return A
   endfunction
end

A_sort = sort(A_vec)
for j=1:length(A_sort)
   CDFA(j) = j/length(A_sort)
end
plot synthetic contour in (p, beta) plane

call function leader_invert
   create (p, beta) grid
   create Mtilde, Ctilde
   W = lsqnonneg(Mtilde, Ctilde)
   return W
endfunction

call function leader_plots
   plot contour in (p, beta) plane
   compute marginal distributions
   draw marginal distributions for p and beta
endfunction

call function leader_postprocess_WISE
   find (p, beta) peak
   dampen bins away from the peak
   P_new = P + deltaP
   draw deconvoluted contour in (p, beta) plane
endfunction

endfunction
```

Table A.2. Pseudocode of the synthetic simulator implementation of LEADER.

```
function ast_comparison_WISE

load population1
call function leader_main_WISE
   return p1, beta1, fp1, fbeta1
endfunction

load population2
call function leader_main_WISE
   return p2, beta2, fp2, fbeta2
endfunction

call function KS_comparison
   create CDFs of fp1, fp2, fbeta1, fbeta2
   % CDFs are called Cp1, Cp2, Cb1, Cb2
   Cp2i = interpolate Cp2 at p1
   Cb2i = interpolate Cb2 at beta1
   for k=[1, 2, inf]
      Dp = alpha_k*||Cp1-Cp2i||_k
      Db = alpha_k*||Cb1-Cb2i||_k
   end
   plot margin distributions in the same figure
   plot CDFs in the same figure
   return Dp, Db
endfunction

endfunction
```

Table A.3. Pseudocode of the statistical comparison of two asteroid populations.

Publication IV

Cibulková, H., Nortunen, H., Ďurech, J., Kaasalainen, M., Vereš, P., Jedicke, R., Wainscoat, R. J., Mommert, M., Trilling, D. E., Schunová-Lilly, E., Magnier, E. A., Waters, C., and Flewelling, H., "Distribution of shape elongations of main belt asteroids derived from Pan-STARRS1 photometry," *Astronomy and Astrophysics*.

A&A 611, A86 (2018)
DOI: 10.1051/0004-6361/201731554

Astronomy & Astrophysics

Distribution of shape elongations of main belt asteroids derived from Pan-STARRS1 photometry

H. Cibulková[1], H. Nortunen[2], J. Ďurech[1], M. Kaasalainen[2], P. Vereš[3], R. Jedicke[4], R. J. Wainscoat[4], M. Mommert[5], D. E. Trilling[5], E. Schunová-Lilly[4], E. A. Magnier[4], C. Waters[4], and H. Flewelling[4]

[1] Institute of Astronomy, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, 180 00 Prague 8, Czech Republic
e-mail: cibulkova@sirrah.troja.mff.cuni.cz
[2] Department of Mathematics, Tampere University of Technology, PO Box 553, 33101 Tampere, Finland
[3] Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA
[4] Institute for Astronomy University of Hawaii, 2680 Woodlawn Dr, Honolulu, HI 96822, USA
[5] Department of Physics and Astronomy, PO Box 6010, Northern Arizona University, Flagstaff, AZ 86011, USA

Received 12 July 2017 / Accepted 10 September 2017

ABSTRACT

Context. A considerable amount of photometric data is produced by surveys such as Pan-STARRS, LONEOS, WISE, or Catalina. These data are a rich source of information about the physical properties of asteroids. There are several possible approaches for using these data. Light curve inversion is a typical method that works with individual asteroids. Our approach in focusing on large groups of asteroids, such as dynamical families and taxonomic classes, is statistical; the data are not sufficient for individual models.
Aims. Our aim is to study the distributions of shape elongation b/a and the spin axis latitude β for various subpopulations of asteroids and to compare our results, based on Pan-STARRS1 survey, with statistics previously carried out using various photometric databases, such as Lowell and WISE.
Methods. We used the LEADER algorithm to compare the b/a and β distributions for various subpopulations of asteroids. The algorithm creates a cumulative distributive function (CDF) of observed brightness variations, and computes the b/a and β distributions with analytical basis functions that yield the observed CDF. A variant of LEADER is used to solve the joint distributions for synthetic populations to test the validity of the method.
Results. When comparing distributions of shape elongation for groups of asteroids with different diameters D, we found that there are no differences for $D < 25$ km. We also constructed distributions for asteroids with different rotation periods and revealed that the fastest rotators with $P = 0 - 4$ h are more spheroidal than the population with $P = 4-8$ h.

Key words. minor planets, asteroids: general – methods: statistical – techniques: photometric

1. Introduction

The spin states (rotational periods and directions of the spin axes) and shapes of individual asteroids can be determined from photometric data by light curve inversion (Kaasalainen & Lamberg 2006; Ďurech et al. 2015, and references therein). For these methods, mainly dense photometric data are used because they sample well the rotational period P. The preliminary estimate of P can substantially reduce the computational time required for the determination of unique sidereal rotational period. Up to now, almost a thousand models have been derived using this method and most of these are stored in Database of Asteroids Models from Inversion Techniques (DAMIT; Ďurech et al. 2010).

In Cibulková et al. (2016), we described a different approach that is suitable for photometric data that are sparse in time and produced by all-sky surveys and consist typically of a few measurements per night over ∼10 yr. These data are not suitable for ordinary sparse light curve inversions (Ďurech et al. 2005, 2007). In Cibulková et al. (2016), we used the mean brightness and its dispersion in individual apparitions to derive the ecliptical longitude and latitude of the spin axis and the shape elongation of asteroids from photometric data stored in the Lowell Observatory database (Bowell et al. 2014). Even though the parameters could be determined for individual asteroids, the uncertainties are large and the results are only supposed to be used in a statistical sense. However, this model cannot be used for the photometric data from the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS) because there are not enough measurements covering long enough time intervals.

Another statistical study was carried out by Nortunen et al. (2017) using data from the Wide-field Infrared Survey Explorer (WISE) database[1]. These data also could not be analyzed with the method from Cibulková et al. (2016), however, Nortunen et al. (2017) developed a new model and described physical parameters for subpopulations of asteroids using distribution functions. This method is not meant to invert the shape and spin characteristics of individual light curves; the inversion works only on a population scale, where we consider the shape and spin distributions of a large population. In Nortunen et al. (2017) as well as in this paper, we constructed cumulative distribution functions (CDFs) of the variation of brightness for selected groups of asteroids and studied the inverse problem. The parameters of the model are the shape elongation b/a and

[1] http://irsa.ipac.caltech.edu/Missions/wise.html

the ecliptical latitude β of the spin axis. The advantage of this method is that it can be used even if only few points and one apparition are available for an asteroid. A similar approach was used by Szabó & Kiss (2008) and McNeill et al. (2016).

While in Nortunen et al. (2017) we mainly studied the validity and accuracy of the method and practical applicability on astronomical databases, in this work we applied the model on photometric data from Pan-STARRS1 and performed an analysis focusing on large subpopulations of asteroids using the Latitudes and Elongations of Asteroid Distributions Estimated Rapidly (LEADER) algorithm (Nortunen & Kaasalainen, 2017). The structure of this paper is as follows: in Sect. 2, we briefly describe the model; in Sect. 3, we describe the used data from Pan-STARRS1 sky survey; in Sect. 4, we test the accuracy of the determination of model parameters by simulations on synthetic data; in Sect. 5, we construct distributions of shape elongations b/a and ecliptical latitudes β of the spin axis for some subpopulations of asteroids and analyze the results; and in Sect. 6, we summarize the main results.

2. Model

In our model, we approximate the shape of an asteroid with a simple, biaxial ellipsoid. We denote the semiaxes $a \geq b = c = 1$, and we choose b/a as the parameter that describes the shape elongation of an asteroid. We have $0 < b/a \leq 1$, with a small b/a presenting an elongated body and $b/a = 1$ presenting a sphere. This shape approximation is very coarse, but with a high number of observations ($\propto 10^3$), it portrays statistical tendencies of a population accurately. For realistic shapes, the proportion of highly elongated values $b/a < 0.4$ is negligible, and for most shapes, $b/a > 0.5$. However, for completeness, we include all the values $0 < b/a \leq 1$ in our grid; if the solved b/a distribution contains an unusually high proportion of values below 0.4, it is usually an indicator of error in the solution, caused by noise and/or instabilities.

Our second parameter is the spin co-latitude β, which is defined as the ecliptic polar angle of the spin axis. The connection between β and the aspect angle of the pole is explained in Nortunen et al. (2017). In the ellipsoid model, the values of β are fixed in the interval $[0, \pi/2]$. In other words, there is no way to distinguish whether the spin latitudes are above or below the ecliptic plane in our model. In our convention, $\beta = 0$ indicates that the spin direction is perpendicular to the ecliptic plane, while $\beta = \pi/2$ means the spin is in the ecliptic plane; this was the convention in Nortunen et al. (2017), but it was the opposite in Cibulková et al. (2016). We assume that most orbits are in the ecliptic plane.

Assuming we have the brightness intensities, L, measured with the data given by an asteroid database, we use the brightness variation η as our observable,

$$\eta = \frac{\Delta(L^2)}{\langle L^2 \rangle}. \tag{1}$$

The squared intensities L^2 are used for the mean $\langle L^2 \rangle$ and the variation $\Delta(L^2) := \sqrt{\langle (L^2 - \langle L^2 \rangle)^2 \rangle}$ instead of the standard brightness L to obtain more simple, closed-form formula for η. From Nortunen et al. (2017), the amplitude A can be directly computed from η as follows:

$$A = \sqrt{1 - \left(\frac{1}{\sqrt{8}\eta} + \frac{1}{2} \right)^{-1}}. \tag{2}$$

The amplitude A is based on intensity here, not on magnitudes. With the amplitudes known, we can create their CDF, $C(A)$.

To solve the joint distribution for b/a and β, we create a grid of bins $((b/a)_i, \beta_j) \in [0, 1] \times [0, \pi/2]$, where i = 1, ..., k and $j = 1, \ldots, l$. Our goal is to determine the proportion of each bin. Now, the CDF can be written as a linear combination of other functions,

$$C(A) = \sum_{\mathrm{i,j}} w_{ij} F_{ij}(A), \tag{3}$$

where $F_{ij}(A)$ are monotonously increasing basis functions derived by Nortunen et al. (2017), i.e.,

$$F_{ij}(A) = \begin{cases} 0, & A \leq (b/a)_i \\ \dfrac{\pi}{2} - \arccos \dfrac{\sqrt{A^2 - (b/a)_i^2}}{\sin\beta_j \sqrt{1 - (b/a)_i^2}}, & (b/a)_i < A < \mathcal{F}((b/a)_i, \beta_j) \\ \dfrac{\pi}{2}, & A \geq \mathcal{F}((b/a)_i, \beta_j), \end{cases} \tag{4}$$

where $\mathcal{F}((b/a)_i, \beta_j) = \sqrt{\sin^2\beta_j + (b/a)_i^2 \cos^2\beta_j}$. Each basis function $F_{ij}(A)$ describes the contribution made by objects in a given bin $((b/a)_i, \beta_j)$ to the CDF $C(A)$. The weights w_{ij} are the occupation numbers of each bin $((b/a)_i, \beta_j)$. We can write (3) in an equivalent form,

$$Mw = C, \tag{5}$$

where each column of the matrix M contains a basis function $F_{ij}(A)$, the vector w contains the occupation numbers w_{ij}, and the vector C contains the CDF $C(A)$. For solving (5), we use linear least squares methods in, for example, Matlab, along with regularization and a positivity constraint that $w_{ij} \geq 0$. With the weights w_{ij} solved, we have the proportion of each bin $((b/a)_i, \beta_j)$.

With the joint distribution for b/a and β obtained, we compute the marginal DFs $f_{b/a}$ and f_β for both parameters,

$$f_{(b/a)_i} = \sum_{j=1}^{l} w_{ij}, \quad f_{\beta_j} = \sum_{i=1}^{k} w_{ij}. \tag{6}$$

In addition, we compute the CDFs for the marginal DFs. Let us denote the CDFs as $F_{b/a}$ and F_β. Now, we assume that we obtained these CDFs for two subpopulations, S_1 and S_2, where the CDFs are denoted as $F_{b/a}(S_1)$, $F_{b/a}(S_2)$, $F_\beta(S_1)$, and $F_\beta(S_2)$, and we want to measure statistical differences of the populations. Some of these measures were used in Nortunen et al. (2017) as follows:

$$D_{b/a}(S_1, S_2) = \alpha_k \left\| F_{b/a}(S_1) - F_{b/a}(S_2) \right\|_k, \tag{7}$$

$$D_\beta(S_1, S_2) = \alpha_k \left\| F_\beta(S_1) - F_\beta(S_2) \right\|_k, \tag{8}$$

where $k = 1; 2; \infty$ and α_k are norm-based scaling factors, $\alpha_1 = 1/4$; $\alpha_2 = 1$ and $\alpha_\infty = 2$. Each norm provides a different kind of information about the statistical differences of the populations. The case $k = \infty$ corresponds with the Kolmogorov-Smirnov test (for details see Nortunen et al. 2017; or Nortunen & Kaasalainen 2017). As a general rule of thumb, two distributions are considered significantly different statistically if $D \gtrsim 0.2$. However, a visual inspection on the marginal DF and CDF plots is also recommended for obtaining a better understanding of the statistical

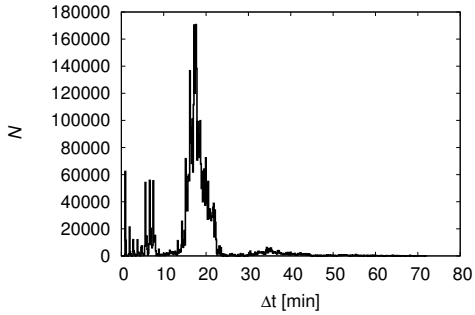

Fig. 1. Histogram of time intervals between measurements in the w-band filter from the Pan-STARRS1 survey.

differences. The detailed description of the LEADER software can be found in Nortunen & Kaasalainen (2017) and the software itself is available in DAMIT database[2].

3. Data

The 1.8 m Pan-STARRS1 survey telescope (Hodapp et al. 2004; Tonry et al. 2012), build atop of Haleakala, Maui, started its three-year science mission in May 2010. Photometric data were obtained in six optical and near-infrared filters (g, r, i, z, y, and w). Because of the distinct survey goals and patterns, most of the asteroids were observed in a wideband w-filter (~400–700 nm). We used the unpublished high-precision calibrated chip-stage photometry (Schlafly et al. 2012) with photometric errors and selected detections of a good photometric quality. Only PSF-like and untrailed detections were considered. Our subset spanned from April 11, 2011 until May 19, 2012. In total, we had photometric data for 348 210 asteroids with about 20 measurements for an asteroid on average. The second highest number of measurements was in the i band, where we had data for 136 463 asteroids. Only the w-band data provided enough measurements for a reasonable application of our model. We briefly discuss the results from the i-filter and compare these with results from the w-filter in Sect. 5.7.

The typical time interval between two measurements in the w-band filter is ~17 min (see Fig. 1). However, not all the data were applicable to our model. Our conditions on the data were the following:

1. The time interval between measurements is greater than 0.01 day (~14 min). In the case of a shorter interval the rotational period would not be randomly sampled over one rotation of ~hours, and in the case of a longer minimum interval we would lose a significant amount of data, as we can see from Fig. 1.
2. Then, we limited the solar phase angle α to be $\leq 20°$. In the model we assume this angle to be close to zero, however, in the data, there are not enough measurements with $\alpha \sim 0°$, therefore, we have to choose some reasonable value (see also Fig. 2). As described in Nortunen et al. (2017; they used $\alpha \leq 30°$) the error caused by this condition is negligible.
3. Finally, we required at least five measurements satisfying previous conditions within three days to keep the geometry of observation sufficiently constant; this is the same condition as in Nortunen et al. (2017).

[2] http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php

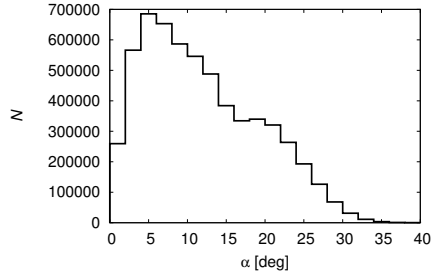

Fig. 2. Histogram of the solar phase angle α for measurements in the w-filter.

It is possible that for the same asteroid we had two (or even more) sets of measurements. In that case, each set was incorporated in the model.

4. Synthetic simulations for accuracy estimation

Before we computed the solution of the inverse problem from Eq. (5) for any Pan-STARRS1 subpopulation, we performed a thorough analysis concerning whether the method is reliable and accurate with the given database. To do this, we used synthetic data created according to the procedure described in Nortunen & Kaasalainen (2017). We chose a peak of the $(b/a, \beta)$ distribution. For each asteroid in the considered population we chose a shape model from DAMIT, with $|b/a_{\mathrm{DAMIT}} - b/a_{\mathrm{wanted}}| \leq 0.075$. The rotation period was chosen randomly between 3 and 12 h from a uniform distribution; we did not use rotation periods from DAMIT, as they could be biased. Next, we used the real Pan-STARRS1 geometries and times of observations and computed the synthetic brightness using a combination of Lommel-Seeliger and Lambert scattering laws. To simulate noise, we added a minor Gaussian perturbation 1–2%. Our aim was to find how well the solution distribution computed from Eq. (5) coincides with the known, synthetic distribution. For simplicity, we were interested in reconstructing the highest peak of the joint $(b/a, \beta)$ distribution. The peak is defined as the bin with the highest occupation numbers. If there were any obvious systematic errors in the computed solution, we attempted to apply a posterior correction to the solution. Similar synthetic simulations were used by Nortunen et al. (2017) to estimate the accuracy of the method for the WISE database and to create a deconvolution filter to the contour image of the solution.

4.1. Number of bodies in a population

We created 50 synthetic populations, each containing N asteroids, and each population having a distinct, single peak chosen randomly. With each population, we plotted the actual $(b/a, \beta)$ peak versus the computed $(b/a, \beta)$ peak to see how well they coincide. We set the populations to have from 100 to 5000 asteroids. This was so that we could evaluate how the accuracy of our method increases with a growing number of asteroids. The results from these simulations were plotted in Fig. 3 for b/a and Fig. 4 for β.

The b/a plots in the left column of Fig. 3 show that the accuracy of the obtained b/a distribution is improved substantially as the population size increases. There is always some overshoot and undershoot when $b/a \lesssim 0.4$, but this is a rare problem; with real data, we typically have $b/a \gtrsim 0.5$, so the peak of

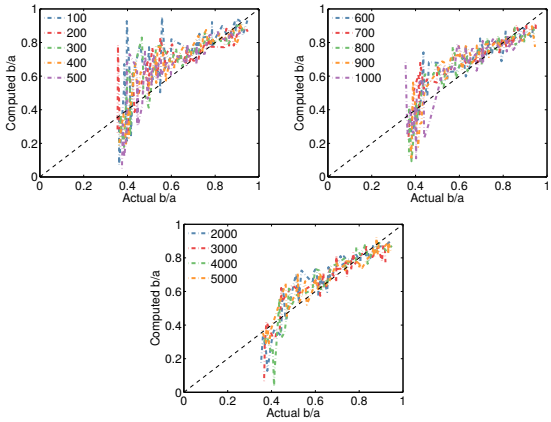

Fig. 3. Synthetic simulations showing how the accuracy of our method improves for b/a with a growing number of asteroids (from 100 to 5000). The plots have the real peak of the distribution plotted vs. the computed peak. The black dashed line of the form $y = x$ depicts the ideal situation when the actual and computed peaks are the same.

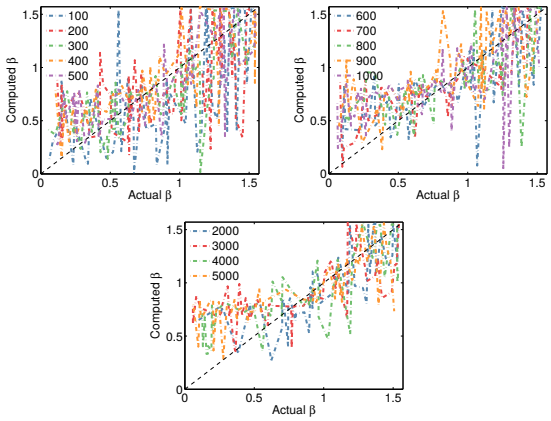

Fig. 4. Synthetic simulations similar to Fig. 3, but for β.

the distribution is also expected to be above 0.5. With a population of less than 1000 asteroids, there is a slight overshoot when $b/a > 0.5$; i.e., the solution suggests the shapes are slightly more spherical than they actually are. But as the population size exceeds 1000 asteroids, the computation of the b/a peak is very accurate when $b/a > 0.5$.

Unfortunately, much of the β information is lost in the inversion carried out for the Pan-STARRS1 database, as seen in Fig. 4. For a population of less than 500 asteroids, no actual information can be recovered. For 600–1000 bodies, there is a slight increase in accuracy, but overall, the solution is too noisy to provide accurate information on β. The improvement of the accuracy is noticeable for populations with 2000–5000 asteroids, and the method provides a rough estimate on the location of the peak; when the β peak is low (perpendicular to the ecliptic plane), the obtained solution also has a low β peak and vice versa. With the Pan-STARRS1 database, our assumption that the majority of the orbits is in the ecliptic plane may not hold well, which considerably reduces the accuracy of the beta distribution.

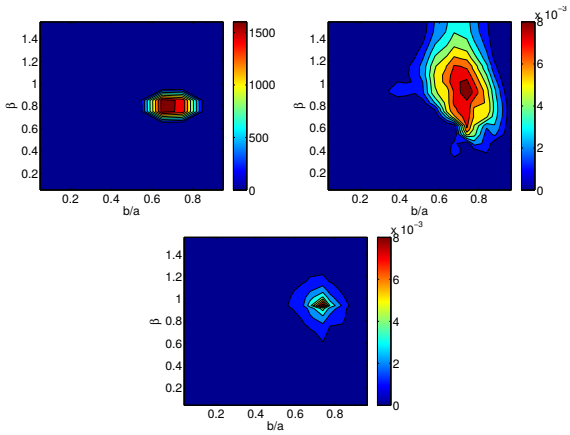

Fig. 5. Synthetic simulations using a fixed $(b/a, \beta)$ peak for a population of 10 000 asteroids. The *top left* plot shows the actual $(b/a, \beta)$ distribution, the *top right* shows the computed $(b/a, \beta)$ distribution, and the *bottom* shows the *top right* solution with a deconvolution filter applied.

Because of the low accuracy of the β solution, we recommend that caution is used when interpreting the computed β distribution. At best, our method can provide a coarse approximation regarding where the β peak is located.

In addition to determining the correct position of the peaks, we are interested in the overall shape of the joint distribution. It is a typical tendency that the computed distribution spreads too much, especially in β direction, and the distribution has a heavy tail toward the spin directions in the ecliptic plane. To correct this error, we applied a deconvolution filter to the computed distribution. In this post-solution correction, we introduced some dampening by reducing the occupation numbers of bins when moving further away from the highest peak, that is, the bin with the biggest occupation number. A similar method was used in Nortunen et al. (2017). An example of a typical solution and the effects of deconvolution is plotted in Fig. 5. In the simulation, we used a single, fixed $(b/a, \beta)$ peak for a population of 10 000 asteroids, with the geometries from Pan-STARRS1 database. We only reduced the spreading of the solved distribution in the post-solution correction; we did not shift the position of the $(b/a, \beta)$ peak.

We emphasize that the accuracy of the solution has a strong dependence on the asteroid database used. Our method should never be used as a black box for a database. Instead, whenever we begin to use a new database, we should always test the validity of our method with synthetic simulations. As the level and distribution of noise in the database is rarely known, synthetic simulations are typically the only way to estimate the error of our method. For comparison, we performed similar synthetic simulations for WISE database in Nortunen & Kaasalainen (2017), and the results obtained from WISE and Pan-STARRS1 databases are considerably different.

4.2. Influence of the rotation period

Next, we studied how accurately we are able to reproduce the known $(b/a, \beta)$ distribution when we created synthetic data with different rotation periods P. We chose the following intervals of P: (i) 3–12 h; (ii) 12–24 h; and (iii) 24–96 h. The syn-

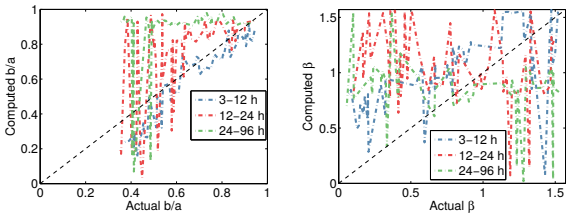

Fig. 6. Synthetic simulations showing the accuracy of our method for different values of rotation period P. The black dashed line denotes the ideal situation.

thetic populations contained 2000 asteroids each. The results are plotted in Fig. 6. Considering the b/a distribution, for $P < 12$ h our method provides reliable results. For $P > 12$ h the solution prefers values of $b/a \sim 1$ (spheroidal bodies) and moreover, the solution becomes unstable for $b/a < 0.6$. As to the β distribution, for $3 < P < 12$ h we noticed a correlation between actual and computed β, but for $P > 12$ h, the β is too unstable to recover any accurate information about the distribution.

The fact, that our computed distributions of b/a for slow rotators ($P > 12$ h) peak at $b/a \sim 1$ is probably due to the time distribution of Pan-STARRS1 measurements. For most asteroids, data were obtained during a single night, i.e., a few hours. If the real P is much longer, the data cover only a small fraction of the full light curve (showing the time evolution of brightness during the whole P). The changes of brightness are thus small and our model interprets these changes as belonging to a spheroidal asteroid. We constructed the distribution of P from the Asteroid Lightcurve Database[3] (LCDB; Warner et al. 2009) for the asteroid included in Pan-STARRS1 database and we found that most of the asteroids have $P \lesssim 15$ h. Nevertheless, the sample of objects in the LCDB database is biased and the number of slow rotators is underestimated since it is observationally difficult to determine long periods (Marciniak et al. 2015; Szabó et al. 2016).

4.3. Influence of the orbit inclination

Finally, we tested the influence of the orbit inclination $\sin I$ on our solution since in the model we assume $\sin I = 0$. When creating the synthetic data, we used Pan-STARRS1 geometries of 2000 asteroids with $\sin I \leq 0.2$, i.e., the first population, and 2000 asteroids with $\sin I > 0.2$, i.e., the second population. The resulting distributions of b/a and β are not statistically different for populations with small and high inclinations of orbits. For b/a, the computed peak corresponds with the actual peak, but for β, we find that the model shifts the peak to middle values, which is the same problem as in Fig. 4.

5. Distributions of the ratio of axes *b/a*

In this section, we first test how many asteroids have to be in a studied subpopulation to obtain reliable results because typically we compare subpopulations that contain different numbers of asteroids. Then we construct the distributions of shape elongation b/a for various subpopulations of main-belt asteroids. Specifically, we tested asteroids with different diameters, different rotation periods, dynamical families, taxonomic classes, and subpopulations of asteroids located in various parts of the main belt

[3] http://www.minorplanet.info/lightcurvedatabase.html

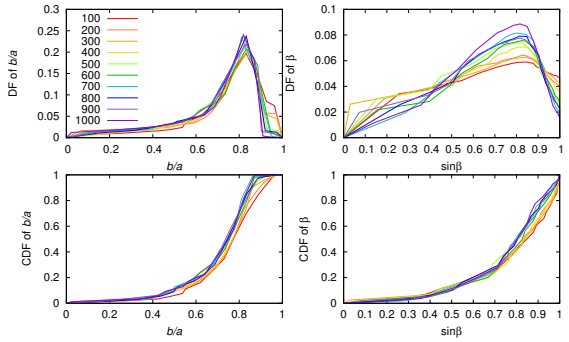

Fig. 7. Distributions of b/a and β for Flora family constructed for growing number of asteroids that were included (from 100 to 1000).

(as in Cibulková et al. 2016). To compare the distributions, we calculated $D_{b/a}$ and D_β according to Eqs. (7) and (8). The bins in the distributions of b/a and β were chosen randomly, hence, for each two subpopulations that were compared, we processed 10 runs and obtained 10 values of $D_{b/a}$ and D_β, from which we calculated the mean values. For the distribution of b/a we chose 14 bins from 0 to 1, however, because the shape elongation $b/a < 0.25$ is improbable, there was only one bin from 0 to 0.25, then one bin from 0.25 to 0.4 and 12 bins from 0.4 to 1. For the distribution of β we chose 20 bins from 0 to $\pi/2$, specifically, 15 bins for $\beta > 43.4°$ and then we always selected one bin in following intervals: 37.2°−43.4°, 31°−37.2°, 24.7°−31°, 18.5°−24.7°, and 0°−18.5° to consider that the distribution pole latitudes is uniform in $\sin\beta$.

5.1. Effect of the number of asteroids in a subpopulation

When comparing subpopulations with each other we have to take into account that they contain different numbers of asteroids. To find which population is large enough for stable results, we performed the following test. We used data for the Flora family and we randomly chose 100 of its members and ran our model 10 times. We obtained 10 distributions of b/a and β from which we calculated one mean distribution of b/a and one for β. We repeated this for a sample of 200 randomly chosen asteroids, then 300, 400, and so on, up to the sample of 2000 asteroids. All mean distributions of these subpopulations of Flora are shown in Figs. 7 and 8. For the distribution of b/a we see that the results are stable from ~800 asteroids in the subpopulation. However, for β the results are much more unstable, the distributions are clearly different even in Fig. 8 that contains populations with 1100 to 2000 asteroids. With a growing number of bodies the peak of β distribution is higher and the number of asteroids with $\beta \sim \pi/2$ decreases.

5.2. Asteroids with different diameters

First, we focused on groups of asteroids with different diameters D. For asteroids that have D derived from the observations of WISE satellite, we used that value and for other asteroids we used diameters from the AstOrb catalog. We divided asteroids into seven groups with $D < 3$ km, 3−6 km, 6−9 km, 9−12 km, 12−15 km, 15−25 km, and $D > 25$ km, and compared these groups with each other. For $D < 15$ km, we have in all five groups more than 1200 asteroids, however, there are only

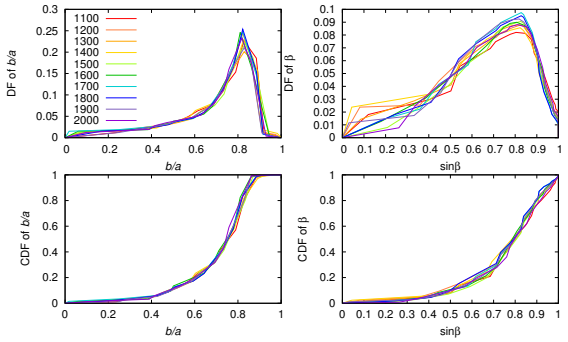

Fig. 8. Same as Fig. 7, but for greater number of asteroids (from 1100 to 2000).

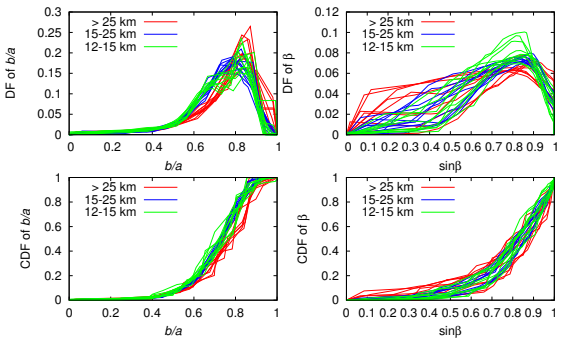

Fig. 9. DFs and CDFs of b/a and β for asteroids with $12 < D < 15$ km (red lines), $15 < D < 25$ (blue lines) and $D > 25$ km (green lines).

990 asteroids with $15 < D < 25$ km and only 223 bodies in the last group ($D > 25$ km), which is not enough for a reliable result. The distributions for three groups with the largest D are shown in Fig. 9. Although, this is in agreement with the findings of Cibulková et al. (2016) that the asteroids larger than $D > 25$ km are more often spheroidal, in this case it might be just an effect of the low number of asteroids in the last subpopulation.

The mean values of $D_{b/a}$ for distributions of b/a for the three subpopulations with largest diameters are listed in Table 1. The distributions of b/a for groups of asteroids with $D < 15$ km are not statistically different from the group of asteroids with $15 < D < 25$ km and have a maximum for $b/a \sim 0.8$. The average axial ratio b/a from Pan-STARRS1 survey was also determined by McNeill et al. (2016). For asteroids with $D < 8$ km these authors found the average b/a to be 0.85, which is a little more spheroidal than our result. For $D < 25$ km, Cibulková et al. (2016) found the maximum of distribution of b/a for ~0.63, i.e., the asteroids are more elongated, nevertheless, the possibility is mentioned there that the results could be influenced by the underestimated data noise, which causes shape estimates to be more elongated.

We also tried to reconstruct the cumulative distributions of absolute rate of change in magnitude from work McNeill et al. (2016), who constructed distributions for asteroids with $1 < D < 8$ km dividing them into groups 1−2 km, 2−3 km, and so on, to 7−8 km. They found that with decreasing diameter, the distributions show smaller change in magnitude. However, we could not find any differences between individual distributions (see Fig. 10 on the left). The possible explanation of this disagreement is that McNeill et al. (2016) used only measurements with magnitude uncertainty ≤0.02, however, we used all measurements; our only conditions were first, the solar phase angle $\alpha < 10°$, which is the same condition as in McNeill et al. (2016), and, second, pairs of measurements separated by time interval 10 min $< \Delta t <$ 20 min. We also constructed cumulative distributions of brightness variation η to see whether there would be any differences, but as shown in Fig. 10 on the right, the η distributions for groups of asteroids with different diameters are almost the same.

Then we focused on the distributions of β. As we can see in Fig. 9 on the right, they look different from results of, for example, Cibulková et al. (2016) or Hanuš et al. (2011), where β is clustered around 0 due to the Yarkovsky-O'Keefe-Radzievskii-Paddack (YORP) effect, which shifts β near the pole of the ecliptic (e.g., Pravec & Harris 2000; Rubincam 2000). Nevertheless, as explained in Sects. 4 and 5.1, we found that the distribution of β is considerably influenced by the number of asteroids in a given subpopulation and becomes flatter with decreasing number of asteroids. In Fig. 4 we also see that the model tends to shift the peak to the middle values. The results on β are thus not reliable and in the following tests we only focus on the distributions of b/a.

Table 1. Parameter $D_{b/a}$ for selected pairs of populations that were compared.

Populations	$D_{b/a}(L^1)$	$D_{b/a}(L^2)$	$D_{b/a}(L^\infty)$
D = 15−25 km; >25 km	0.164	0.269	0.351
D = 12−15 km; 15−25 km	0.091	0.146	0.189
P = 0−4 h; 4−8 h	0.369	0.537	0.573
P = 0−4 h; 8−15 h	0.107	0.163	0.204
P = 4−8 h; 8−15 h	0.450	0.638	0.642
Flora; background	0.087	0.159	0.244
Massalia; background	0.294	0.462	0.554
Nysa Polana; background	0.140	0.259	0.399
Vesta; background	0.068	0.114	0.170
Phocaea; background	0.175	0.274	0.367
Eunomia; background	0.079	0.123	0.176
Gefion; background	0.132	0.203	0.270
Maria; background	0.078	0.129	0.186
Koronis; background	0.142	0.244	0.367
Eos; background	0.084	0.142	0.208
Hygiea; background	0.098	0.163	0.218
Themis; background	0.134	0.243	0.342
Alauda; background	0.144	0.219	0.250
C class; S class	0.081	0.129	0.174
Massalia; background (*i*-filter)	0.273	0.415	0.495
Phocaea; background (*i*-filter)	0.212	0.291	0.333
w-filter; *i*-filter for Nysa Polana	0.095	0.160	0.220

Notes. The given values are the mean values from 10 runs of our model.

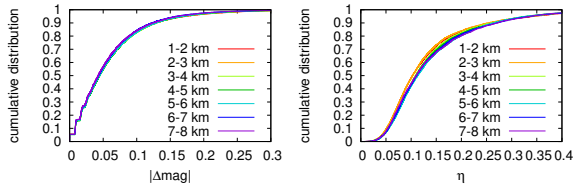

Fig. 10. *Left*: cumulative distributions of absolute value of change in magnitude |Δmag| for groups of asteroids with different sizes. *Right*: cumulative distributions of η for groups of asteroids with different sizes.

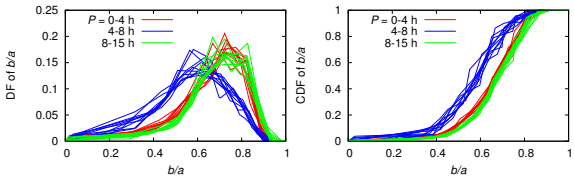

Fig. 11. DFs and CDFs of b/a for asteroid populations with different rotation periods.

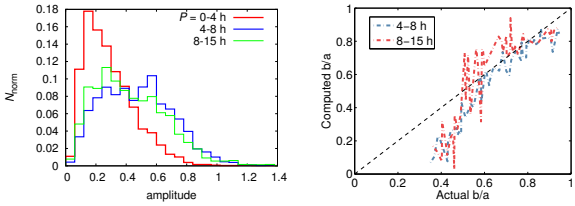

Fig. 12. *Left*: distributions of light curve amplitudes from the LCDB database for different rotation periods P. *Right*: synthetic simulations showing the accuracy of our method for two different intervals of rotation period P.

Because the number of asteroids with $D > 25$ km in data from Pan-STARRS1 is insignificant in comparison to the number of smaller asteroids (less than 1%), this dependence on diameter does not influence the results of the following tests.

5.3. Different rotation periods

According to their rotation periods P provided by the LCDB database, we divided asteroids into three groups. To ensure that all groups are populous enough for stable results we chose the following intervals: (i) $P = 0{-}4$ h (1081 bodies); (ii) $4{-}8$ h (1967 bodies); and (iii) $8{-}15$ h (1071 bodies). We excluded asteroids with $P > 15$ h since our simulations with synthetic data showed the results are not reliable (see also Fig. 6).

We compare populations with each other and plot their distributions of b/a in Fig. 11. We see that the fastest rotators ($P = 0{-}4$ h) are on average more spheroidal than the population with $P = 4{-}8$ h, but their b/a distribution is not different from the third population with $P = 8{-}15$ h. The mean values of $D_{b/a}$ are listed in Table 1.

The critical rotation rate is, for the same density, dependent on the elongation (Pravec & Harris 2000). The spheroidal bodies are thus able to rotate faster that the elongated bodies, which is in accordance with our results for the first two populations. However, we were not able to explain why the third population, with $P = 8{-}15$ h, should contain more spheroidal asteroids than the population with $P = 4{-}8$ h. Therefore, using the LCDB database we constructed distributions of light curve amplitudes for the three above-mentioned populations (see Fig. 12). Higher amplitudes correspond to larger elongations. The distributions of the first two groups are in accordance with the results from Pan-STARRS1 data, but for the third population ($P = 8{-}15$ h), we obtained similar distribution as for the population with $P = 4{-}8$ h.

To explain this discrepancy we performed another test with synthetic data. We used the same setup as in Sect. 4.2, where we studied the influence of the rotation period on the accuracy of the solution, but we chose populations with $P = 4{-}8$ h and $P = 8{-}15$ h. The resulting distributions of b/a are shown in Fig. 12, on the right. The figure shows that our model is not able to correctly reproduce peak $b/a \lesssim 0.6$ for either population, nevertheless such elongation peak is uncommon; most of the asteroids have $b/a > 0.6$. Considering $0.8 > b/a > 0.6$, for the population with $P = 8{-}15$ h, our model provides slightly more spheroidal objects ($b/a \sim b/a_{\rm actual} + 0.1$) and for the population with $P = 4{-}8$ h, it provides slightly more elongated objects ($b/a \sim b/a_{\rm actual} - 0.05$). We conclude that the difference between b/a distributions for these two populations (shown in Fig. 11) is due to the method bias that shifts their b/a values ~0.15 apart.

5.4. Period from estimated photometric slopes

From this analysis we learned that our distributions of b/a for other asteroid populations can be strongly influenced by the appropriate period distributions. Unfortunately, our model does not provide the rotation period P and the LCDB database contains P for only ~14 000 asteroids. That sample, divided into individual populations, is not large enough for a statistical purpose. Nevertheless, if there are many measurements for an asteroid and if they are appropriately distributed in time P can be formally calculated directly from photometric data. More precisely, we need pairs of measurements close in time and also a sufficient number of such pairs.

First we derived a general result for the time series of any signal I that is of pure sinusoidal form of nth order only, augmented by the mean term I_0 (0th order),

$$I = I_0 + \cos n\omega t, \tag{9}$$

where ω is the rotation frequency; we chose this form since the starting point is irrelevant. If the estimates of the time derivative $\mathrm{d}I/\mathrm{d}t$ are available (i.e., measurements of I within a short time interval as with Pan-STARRS1), we can use these to estimate ω and hence the period $P = 2\pi/\omega$ in a simple manner. Using the variation (standard deviation) Δ as defined with Eq. (1), with $I = L^2$, and computing the mean $\langle|\mathrm{d}I/\mathrm{d}t|\rangle$ from Eq. (9) by integrating $\mathrm{d}I/\mathrm{d}t$ over the interval $[0, \pi/2]$, we directly obtain

$$P = 4n\sqrt{2}\Delta(I)/\langle|\mathrm{d}I/\mathrm{d}t|\rangle. \tag{10}$$

Since $I = L^2$ for an ellipsoid is of the pure $n = 2$ double-sinusoidal form (Nortunen et al. 2017), we used Pan-STARRS1 slope estimates $|\mathrm{d}L^2/\mathrm{d}t|$ and their mean $\langle|\mathrm{d}L^2/\mathrm{d}t|\rangle$ to obtain the period with the aid of Eq. (10). However, for each asteroid this requires a number of slope estimates. The derivation $\mathrm{d}L^2/\mathrm{d}t$ can be approximately calculated from pairs of measurements that are close in time, but there is a lower limit due to the accuracy of data. We chose $\mathrm{d}t > 10$ min to distinguish the change of brightness from data noise.

To verify if this relation can be used in practice, we performed the following test on synthetic data. We used the DAMIT models and the Hapke scattering model (Hapke 1981, 1993) with randomly chosen parameters and randomly chosen rotational period P (uniformly distributed from 2 to 50 h). Next, we calculated synthetic brightness,which we assigned to ~1000 asteroids observed with Pan-STARRS1 (leaving the geometry of observations unchanged), for which we had the largest number of measurements. From these new synthetic brightnesses we calculated the period P according to Eq. (10) that should approximate the synthetic P.

The derivative $\langle|\mathrm{d}L^2/\mathrm{d}t|\rangle$ was computed from pairs of measurements separated by time interval $10 < \Delta t < 20$ min and we required at least 12 pairs (to calculate the mean value) within five days. The variation ΔL^2 was also computed within five days.

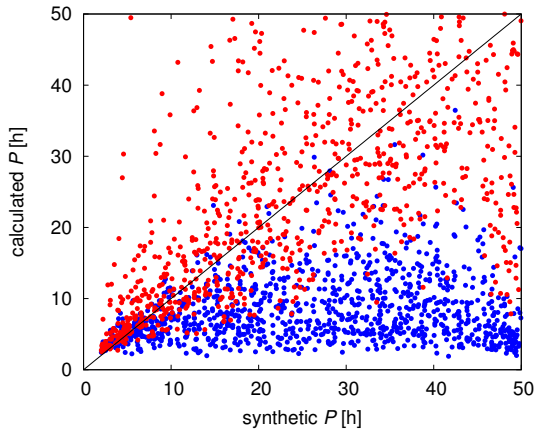

Fig. 13. Comparison of calculated and synthetic rotational period P. The red points denote synthetic data without noise, the blue points synthetic data with noise 0.02.

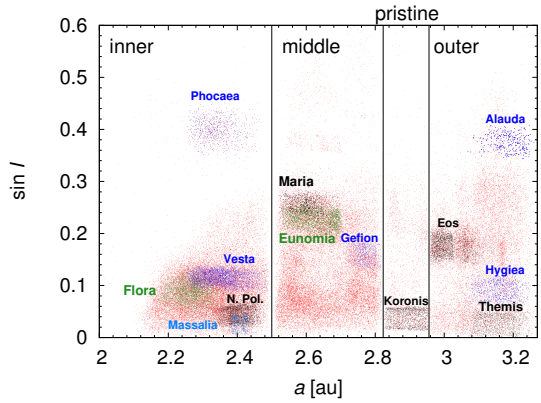

Fig. 14. Four parts of the main asteroid belt defined according to the proper semimajor axis a; we used proper values of a and I from Asteroids Dynamic Site (Knežević & Milani 2003).

We tested synthetic data without any noise and also data with Gaussian noise of 2%. We compared the calculated P with the synthetic P by computing the correlation coefficient; data with noise show no correlation (the coefficient is 0.19) and Fig. 13 (blue points) shows there is a strong preference for low values of P. Interestingly enough, the bias is systematic and amounts to an underestimation factor of about 0.5 for the point fan. Apparently noise systematically increases the slope average from the pairwise slope estimates. The situation for data without noise is slightly better (coefficient 0.30) and if we consider only periods from interval 2 to 30 h, the correlation coefficient is 0.65 (see also Fig. 13). For periods under 10 h, the points are even more tightly clustered near the $x = y$ correlation line.

The possible reason for this bad correlation could be the insufficient number of measurements from which the mean values are calculated. Therefore, to each measurement we added two others, one 0.01 d (14.4 min) earlier and the second 0.01 d later. In total, we had three times more measurements for each asteroid. However, the resulting P were not significantly different from the previous test; in the interval of P from 2 to 30 h, the correlation coefficient is 0.60.

We also tested the relation (10) on real data from Pan-STARRS1 survey, however, there were only few asteroids for which we had the required number of measurements (as described above) and at the same time also the information about the real rotational period from the LCDB database. For these bodies we did not obtain a good agreement between the estimated and real periods. Apparently the use of the period estimate Eq. (10) requires a large number of well-distributed slope pairs over a rotation cycle. Also, a low number of pairs exacerbates the effects of noise and deviations from the pure double-sinusoidal form. Estimates based on the derivative of a function are usually considerably more unstable than those based on the function itself. This approach is thus not applicable in practice and we are not able to correct b/a distributions of other asteroid populations to have the same P distributions.

5.5. Dynamical families

Next, we compared distributions of dynamical families with their background. The family membership of asteroids was taken from Nesvorný et al. (2015). The background for a family is formed by asteroids from the same part of the main belt as the family (inner, middle, pristine, and outer), which do not belong to any other family. We focused on 13 of the most populous families: Vesta, Massalia, Flora, Nysa Polana, and Phocaea in the inner belt; Eunomia, Gefion, and Maria in the middle belt; Koronis in the pristine belt; and Themis, Eos, Hygiea and Alauda in the outer belt (see also Fig. 14). The typical number of asteroids (for which we have enough data) in a family is few thousand, for Vesta, Flora and Nysa Polana it is slightly more than ten thousand and for Phocaea and Alauda it is less than 1000 (the exact numbers are in Table 2). Unlike Cibulková et al. (2016), who did not reveal any differences among families, we found that Massalia has a significantly different distribution of b/a from its background, containing more elongated asteroids. Distributions are shown in the left panel of Fig. 15. Also the cumulative distributions of brightness variation η of Massalia and its background are significantly different; these are shown in the left panel of Fig. 16. Unfortunately, we cannot compare our distribution of b/a for Massalia with the distribution from Nortunen et al. (2017) based on WISE data because their sample contained an insufficient number of bodies. The mean values of $D_{b/a}$ for all families are listed in Table 1. The second largest difference between distribution of b/a is for the Phocaea family and its background (see Fig. 15 on the right), nevertheless the value $D_{b/a}(L^1) = 0.175$ is not high enough for a definite answer. For Phocaea we only have data for 812 asteroids, however, the small number of asteroids cause the population to be more spheroidal and, as we can see in Fig. 15 on the right, Phocaea, in comparison to its background, contains more elongated objects.

The difference between Massalia family and its background could be due to the different period distributions. To test this possibility we used the LCDB database and constructed distributions of P for Massalia and its background. We found that Massalia really contains fewer objects with $P = 0$–4 h and more with $P = 4$–8 h than its background, which is in accordance with the family members being more elongated (compare with Fig. 11). However, we have to emphasize that the distribution of P for Massalia contains only 100 bodies and its background is represented by 420 bodies, which is not enough for a solid conclusion. For the Phocaea family, we do not have enough determined periods to perform such test as for Massalia.

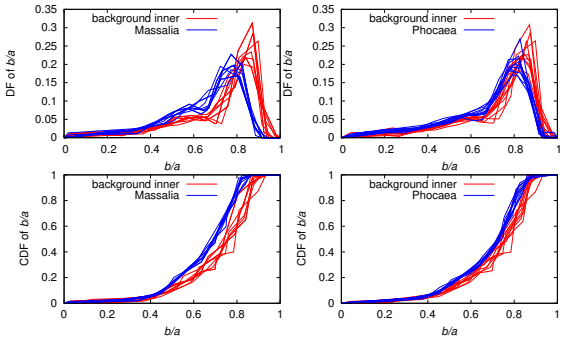

Fig. 15. *Left*: DFs and CDFs of b/a for Massalia family (blue lines) and its background (red lines). *Right*: the same for Phocaea family is shown.

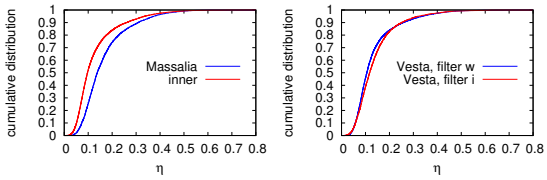

Fig. 16. *Left*: cumulative distributions of brightness variation η for Massalia family and its background. *Right*: the same for Vesta family in filter w and filter i is shown.

Table 2. Number of asteroids in individual families and corresponding backgrounds for which we have data from Pan-STARRS1 survey in filters w and i.

Family	N_w	Background$_w$	N_i	Background$_i$
Flora	11 291	11 029	4135	5316
Massalia	4267	11 029	1032	5316
Nysa Polana	14 741	11 029	4675	5316
Vesta	11 895	11 029	4863	5316
Phocaea	812	11 029	577	5316
Eunomia	4126	12 069	2247	6728
Gefion	2629	12 069	1203	6728
Maria	2203	12 069	1243	6728
Koronis	4845	1272	1881	775
Eos	8237	6665	4272	4172
Hygiea	4191	6665	1584	4172
Themis	4181	6665	1588	4172
Alauda	649	6665	489	4172

Our distributions of b/a look different from the results of Szabó & Kiss (2008), who determined distributions for eight families using data from the Sloan Digital Sky Survey (SDSS); however these authors assumed a fixed value of spin axis latitude for all asteroids, which probably influenced the results. We also did not find any dependence of the distribution of b/a on the age of family that they suggested.

5.6. Taxonomic classes and different parts of the main belt

We also compared the distributions of b/a of the two most populated taxonomic classes: S that dominates in the inner mail belt, and C that dominates in the middle and outer belt. We assigned a taxonomic class to asteroids according to the SDSS-based Asteroid Taxonomy (Hasselmann et al. 2010, data are available on Planetary Data System[4]). For both classes we had data for ∼10 000 asteroids. We did not find these two groups to have different distributions of the shape elongation b/a.

Finally, we compared groups of asteroids with different semimajor axes (inner, middle, pristine, outer) and with different inclinations of orbit. None of the subpopulations is significantly different from others.

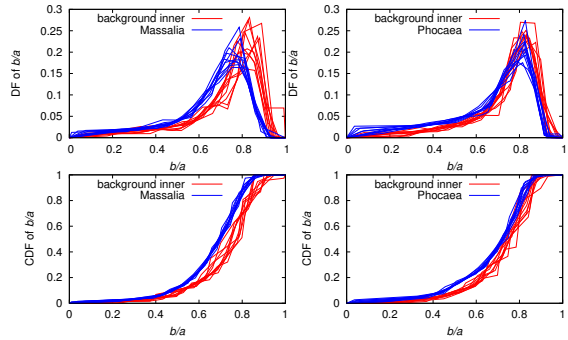

Fig. 17. Same as in Fig. 15, but in filter i.

5.7. Comparison of results from filters w and i

We also analyzed Pan-STARRS1 data in the i-filter (∼700−800 nm, Tonry et al. 2012) and compared the results with the w-filter. We had data for 136 463 asteroids and on average, there were ∼10 measurements for one asteroid. We focused only on taxonomic classes and dynamical families. There were not enough asteroids to study the dependence of the elongation of asteroids on the diameter; only a few asteroids were in the two subpopulations with the largest D.

The number of asteroids in subpopulations containing the taxonomic class S was 6349 and for the taxonomic class C 5813. As in the w-filter, the difference between these two groups is insignificant. Then we focused on dynamical families. As in the w-filter, we found that Massalia family has a significantly different distribution of b/a from its background. Moreover, the result for Phocaea ($D_{b/a}(L^1) = 0.212$) also suggests that this family could have a different distribution of b/a from its background. However, Fig. 17 does not show a significant difference.

To compare results from the filters w and i directly, we constructed distributions of b/a for some families in both filters and calculated $D_{b/a}$. We did not find any significant differences between filters. As an example, distributions of Nysa Polana are shown in Fig. 18. We also constructed cumulative distributions of the brightness variation η for some families in both filters to check that there are no differences between filters before the inversion (see Fig. 16 on the right).

6. Conclusions

In this work, we analyzed photometric data from the Pan-STARRS1 survey via a statistical approach based on cumulative distribution functions. We applied the model and software package LEADER from Nortunen et al. (2017) and Nortunen & Kaasalainen (2017), which allows us to construct distribution functions of the shape elongation b/a and the ecliptical latitude β of the spin axis for some subpopulations of

[4] https://sbn.psi.edu/pds/resource/sdsstax.html

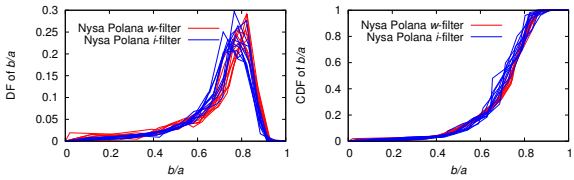

Fig. 18. DFs (*left*) and CDFs (*right*) of b/a for Nysa Polana family in filter i (blue lines) and filter w (red lines).

asteroids and compare these functions with each other. Limitations of this model are that it does not provide the pole longitude and it provides only the combined distribution of the β of both ecliptic hemispheres. Moreover, by testing on synthetic data we found that our model shifts the peak of the β distribution to the middle values and is strongly influenced by the number of objects in studied subpopulations. Distribution of β also appears to be highly sensitive to the used database. For the distribution of b/a we found that the model provides stable results for numbers of objects higher than ~800. The test with synthetic data also revealed that our model provides reliable results only for asteroids with rotation periods $P \lesssim 12$ h. This is due to the time distribution of measurements of the Pan-STARRS1 survey and thus it is not a limitation of the method in general.

We mainly analyzed data in the wide w-band filter. The most populous subpopulations were studied also in the i-filter. The main results of this paper are the following. Groups of asteroids with diameter $D < 25$ km do not have significantly different distributions of b/a; the maximum of these distributions is for $b/a \simeq 0.8$. The distribution for asteroids larger than 25 km suggests that these objects are more spheroidal in comparison with the smaller objects, nevertheless, the number of objects in this subpopulation is insufficient for a strong result. By comparing distributions of b/a for different intervals of rotation period P we found that the fastest rotators with $P = 0$–4 h are more spheroidal (the maximum is for $b/a \sim 0.75$) than the population with $P = 4$–8 h (the maximum is for $b/a \sim 0.6$). We also constructed distributions of b/a for 13 most populous dynamical families and we revealed two families in the inner belt, Massalia and Phocaea, to be significantly different from their background. Both families have members that are more elongated than corresponding backgrounds. One possible explanation is that such a result is due to the dependence of shape elongation on the rotation period. Finally, by analyzing data in the i-filter, we confirmed previous results and we did not find any significant differences between subpopulations studied in the w-filter in comparison with the i-filter.

Acknowledgements. H. Cibulková and J. Ďurech were supported by the grant 15-04816S of the Czech Science Foundation. The research by H. Nortunen and M. Kaasalainen was supported by the Academy of Finland (Centre of Excellence in Inverse Problems), and H.N. was additionally supported by the grant of Jenny and Antti Wihuri Foundation. We would like to thank Matti Viikinkoski for valuable comments and feedback, and assistance with the software.

References

Bowell, E., Oszkiewicz, D. A., Wasserman, L. H., et al. 2014, Meteorit. Planet. Sci., 49, 95
Carvano, J. M., Hasselmann, P. H., Lazzaro, D., & Mothé-Diniz, T. 2010, A&A, 510, A43
Cibulková, H., Ďurech, J., Vokrouhlický, D., Kaasalainen, M., & Oszkiewicz, D. A. 2016, A&A, 596, A57
Ďurech, J., Grav, T., Jedicke, R., Denneau, L., & Kaasalainen, M. 2005, Earth Moon and Planets, 97, 179
Ďurech, J., Scheirich, P., Kaasalainen, M., et al. 2007, in IAU Symp. 236, eds. G. B. Valsecchi, D. Vokrouhlický, & A. Milani, 191
Ďurech, J., Sidorin, V., & Kaasalainen, M. 2010, A&A, 513, A46
Ďurech, J., Carry, B., Delbo, M., Kaasalainen, M., & Viikinkoski, M. 2015, in Asteroid Models from Multiple Data Sources, eds. P. Michel, F. E. DeMeo, & W. F. Bottke, 183
Hanuš, J., Ďurech, J., Brož, M., et al. 2011, A&A, 530, A134
Hapke, B. 1981, J. Geophys. Res., 86, 4571
Hapke, B. 1993, Theory of reflectance and emittance spectroscopy (UK: Cambridge University Press)
Hodapp, K. W., Kaiser, N., Aussel, H., et al. 2004, Astron. Nachr., 325, 636
Kaasalainen, M., & Lamberg, L. 2006, Inverse Problems, 22, 749
Knežević, Z., & Milani, A. 2003, A&A, 403, 1165
Marciniak, A., Pilcher, F., Oszkiewicz, D., et al. 2015, Planet. Space Sci., 118, 256
McNeill, A., Fitzsimmons, A., Jedicke, R., et al. 2016, MNRAS, 459, 2964
Nesvorný, D., Brož, M., & Carruba, V. 2015, in Identification and Dynamical Properties of Asteroid Families, eds. P. Michel, F. E. DeMeo, & W. F. Bottke, 297
Nortunen, H., & Kaasalainen, M. 2017, A&A, 608, A91
Nortunen, H., Kaasalainen, M., Ďurech, J., et al. 2017, A&A, 601, A139
Pravec, P., & Harris, A. W. 2000, Icarus, 148, 12
Rubincam, D. P. 2000, Icarus, 148, 2
Schlafly, E. F., Finkbeiner, D. P., Jurić, M., et al. 2012, ApJ, 756, 158
Szabó, G. M., & Kiss, L. L. 2008, Icarus, 196, 135
Szabó, R., Pál, A., Sárneczky, K., et al. 2016, A&A, 596, A40
Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, ApJ, 750, 99
Warner, B. D., Harris, A. W., & Pravec, P. 2009, Icarus, 202, 134

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-4164-3
ISSN 1459-2045